Efficient Methods for Empirical Tests of Behavioural Economics Theories in Laboratory and Field Experiments.

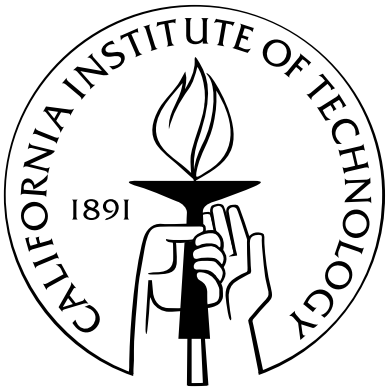
Thesis by

Debajyoti Ray

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2015

(Defended 28 May 2014)

Contents

List of Figures

List of Tables

Acknowledgements

When I started my doctoral program, I was motivated to find the right answers upon completing my thesis. I naively believed that somehow I'd devise methods to query human nature and find interesting answers. As I write my thesis, I realize I have so many more questions to ask and very few right answers. What I have gained during my doctoral thesis years however, is the ability to ask good questions, seek answers with the scientific method, and an appreciation for empirical data. I have been very fortunate to have had some amazing mentors along the way.

During my undergraduate years, I got a taste of machine learning research and the scientific method from Professors Richard Zemel, Radford Neal and Quaid Morris. During my MPhil, Professor Peter Dayan helped me hone my thinking, and ask the right empirical questions.

At Caltech, I learned a lot from Professors Andreas Krause, Peter Bossaerts and Matthew Shum. My advisor, Professor Colin Camerer, made these five years of apprenticeship a pleasurable pursuit of knowledge. I've been infected by his natural intellectual curiosity, and the pursuit of answers for its own sake.

I'm grateful to Brad Feldman, and Pasadena Labs, for providing valuable data for chapter 3 of my thesis. Brad has been a great resource and help me think through the practical aspects of my research. I thank my committee members, Katrina Ligett and Romann Weber, for providing valuable feedback on this thesis.

I'm grateful for my friends and family in Los Angeles, and all over the world, for providing some wonderful times. I'm thankful to Ilona, for this journey wouldn't be as successful or enjoyable without her. Finally, my deepest gratitude is reserved for my parents who instilled in me the values of self-expression and intellectual curiosity, and always supported me.

Abstract

In the quest for a descriptive theory of decision-making, the rational actor model in economics imposes rather unrealistic expectations and abilities on human decision makers. The further we move from idealized scenarios, such as perfectly competitive markets, and ambitiously extend the reach of the theory to describe everyday decision making situations, the less sense these assumptions make. Behavioural economics has instead proposed models based on assumptions that are more psychologically realistic, with the aim of gaining more precision and descriptive power. Increased psychological realism, however, comes at the cost of a greater number of parameters and model complexity. Now there are a plethora of models, based on different assumptions, applicable in differing contextual settings, and selecting the right model to use tends to be an ad-hoc process. In this thesis, we develop optimal experimental design methods and evaluate different behavioral theories against evidence from lab and field experiments.

We look at evidence from controlled laboratory experiments. Subjects are presented with choices between monetary gambles or lotteries. Different decision-making theories evaluate the choices differently and would make distinct predictions about the subjects' choices. Theories whose predictions are inconsistent with the actual choices can be systematically eliminated. Behavioural theories can have multiple parameters requiring complex experimental designs with a very large number of possible choice tests. This imposes computational and economic constraints on using classical experimental design methods. We develop a methodology of adaptive tests: Bayesian Rapid Optimal Adaptive Designs (BROAD) that sequentially chooses the "most informative" test at each stage, and based on the response updates its posterior beliefs over the theories, which informs the next most informative test to run. BROAD utilizes the Equivalent Class Edge Cutting (EC^2) criteria to select tests. We prove that the EC^2 criteria is adaptively submodular, which allows us to prove theoretical guarantees against the Bayes-optimal testing sequence even in the presence of noisy responses. In simulated ground-truth experiments, we find that the EC^2 criteria recovers the true hypotheses with significantly fewer tests than more widely used criteria such as Information Gain and Generalized Binary Search. We show, theoretically as well as experimentally, that surprisingly these popular criteria can perform poorly in the presence of noise, or subject errors. Furthermore, we use the adaptive submodular property of EC^2 to implement an accelerated greedy version of BROAD which leads to orders of magnitude speedup over other methods.

We use BROAD to perform two experiments. First, we compare the main classes of theories for decision-

making under risk, namely: expected value, prospect theory, constant relative risk aversion (CRRA) and moments models. Subjects are given an initial endowment, and sequentially presented choices between two lotteries, with the possibility of losses. The lotteries are selected using BROAD, and 57 subjects from Caltech and UCLA are incentivized by randomly realizing one of the lotteries chosen. Aggregate posterior probabilities over the theories show limited evidence in favour of CRRA and moments' models. Classifying the subjects into types showed that most subjects are described by prospect theory, followed by expected value. Adaptive experimental design raises the possibility that subjects could engage in strategic manipulation, i.e. subjects could mask their true preferences and choose differently in order to obtain more favourable tests in later rounds thereby increasing their payoffs. We pay close attention to this problem; strategic manipulation is ruled out since it is infeasible in practice, and also since we do not find any signatures of it in our data.

In the second experiment, we compare the main theories of time preference: exponential discounting, hyperbolic discounting, "present bias" models: quasi-hyperbolic (α, β) discounting and fixed cost discounting, and generalized-hyperbolic discounting. 40 subjects from UCLA were given choices between 2 options: a smaller but more immediate payoff versus a larger but later payoff. We found very limited evidence for present bias models and hyperbolic discounting, and most subjects were classified as generalized hyperbolic discounting types, followed by exponential discounting.

In these models the passage of time is linear. We instead consider a psychological model where the perception of time is subjective. We prove that when the biological (subjective) time is positively dependent, it gives rise to hyperbolic discounting and temporal choice inconsistency.

We also test the predictions of behavioral theories in the "wild". We pay attention to prospect theory, which emerged as the dominant theory in our lab experiments of risky choice. Loss aversion and reference dependence predicts that consumers will behave in a uniquely distinct way than the standard rational model predicts. Specifically, loss aversion predicts that when an item is being offered at a discount, the demand for it will be greater than that explained by its price elasticity. Even more importantly, when the item is no longer discounted, demand for its close substitute would increase excessively. We tested this prediction using a discrete choice model with loss-averse utility function on data from a large eCommerce retailer. Not only did we identify loss aversion, but we also found that the effect decreased with consumers' experience. We outline the policy implications that consumer loss aversion entails, and strategies for competitive pricing.

In future work, BROAD can be widely applicable for testing different behavioural models, e.g. in social preference and game theory, and in different contextual settings. Additional measurements beyond choice data, including biological measurements such as skin conductance, can be used to more rapidly eliminate hypothesis and speed up model comparison. Discrete choice models also provide a framework for testing behavioural models with field data, and encourage combined lab-field experiments.

Chapter 1

Introduction

The aim of this thesis is to develop rational approaches to study (possibly) irrational behaviour. Human behaviour is studied using different approaches and techniques in several fields, including psychology, sociology, computer science, economics and neuroscience. We'll be investigating behaviour through the lens of economic decision-making.

Economics is concerned with both the *normative* (sometimes *prescriptive*) and *descriptive* aspects of behaviour. How should utility be maximized from limited resources? As well as hypothesis testing of models of decision-making from data. This distinction between how an agent *should* act versus how people *do* act is artificial, from the point of view of neoclassical economics.

The scientific movement of empirically validating economic theories was led by Friedman (1953). Theories make predictions, and hypothesis testing should be utilized to test theories against data. But Friedman (1953) did not question whether the axiomatic foundations or assumptions of the rational actor model, which implies that humans are purely selfish and have unbounded computational abilities, were valid. Instead he proposed that behaviour could be described "as-if" the assumptions held true.

There are at least two compelling arguments to entertain models with different assumptions. This is especially true when decision making models are applied to describe behaviour in everyday situations, and not just perfectly competitive markets. First, the computational complexity makes the application of several solution concepts infeasible for complex tasks. Computing the Nash-equilibria is PPAD-complete in several settings (Daskalakis *et al.* (2009)); even the "consumer problem" of choosing a set of indivisible goods among multiple options with budget constraints is NP-Complete (Gilboa *et al.* (2009)). Second, evidence has been piling up in behavioural economics, that the normative economic models don't always hold up when applied to describe human decision-making and behaviour. At the root, the objection is that the underlying assumptions of *rationality* in these models are unrealistic. Alternate *descriptive* models have been proposed that have more realistic assumptions of human behaviour with roots in Psychology and Cognitive Science.

In developing the necessary tools, the position in this thesis is to follow normative principles, to the extent feasible, of optimization under resource constraints. We develop methods for experimental design to maximize the information gained from conducting tests, or asking questions, with limited time or budgetary resources. We encounter the scourge of computational complexity in experimental design when responses are noisy: the optimal sequence of tests is infeasible to compute. We develop Active-Learning methods inspired from Computer Science and Machine Learning, and prove theoretical guarantees that the resulting solution is *near-optimal*.

When studying the descriptive models, the position in this thesis with respect to the validity of neoclassical and behavioural models is agnostic. We adopt the scientific task of testing several popular models using empirical data from both laboratory and field experiments. In the next section, we provide some background in behavioural and experimental economics. We summarize the key findings and contributions in section 1.4.

1.1 Behavioural Economics

A starting point for Behavioural Economics is querying the explicit, and sometimes implicit, assumptions behind the rational actor framework in the standard model, using lab and field evidence. Based on empirical evidence, behavioural economics increases the explanatory power and precision of economics by providing more realistic psychological foundations (Camerer and Loewenstein (2004)).

The beliefs, values, preferences and choices in the standard model, are built on four main axioms related to consumer preferences: Completeness, Transitivity, Reflexivity and Reveal Preference. Additionally preferences must satisfy dominance[1] and invariance (i.e. the relation of preference should not depend on the description of the options). Furthemore agents have perfect rationality, i.e. they have all the relevant information pertaining to a decision and the cognitive resources to process it instantly and costlessly, and update and estimate their beliefs with Bayesian probability.

Simon (1955) introduced the term *bounded rationality* for the realistic cases where perfect rationality does not hold. The most general implication is that we tend to use *heuristics*, that are "methods for arriving at satisfactory solutions with modest amounts of computation" (Simon (1990)). The most significant implication of using heuristics is that they often result in *biases*, i.e. systematic errors. These are factual errors in beliefs, and also errors in terms of preferences, which may result in non-optimal choice.

Kahneman *et al.* (1982) also documented several studies that showed that people violate rational Bayesian updating. For instance, people overestimate the probability of *salient* events like dying in plane crashes and

[1] Dominance states that if option X is better than option Y in one state and at least as good in all other states, then option X is dominant over option Y and should be chosen. This is related to the reflexivitiy axiom: any bundle is at least as good as itself: $(x_1, x_2) \geq (x_1, x_2)$.

winning the lottery. The representative heuristic refers to the phenomenon that generalized judgement of a category is determined primarily by properties of a prototype (Kahneman and Tversky (1972)); Forming stereotypes is a consequent bias.

Without perfect rationality and Bayesian updating, the revealed preferences approach does not hold up. Behavioural economics, as a field, has documented several anomalies from the standard framework and proposed alternate models to account for these anomalies. Unlike the neoclassical approach, however, a dissatisfying aspect of behavioural economics (in its current state) is that it does not have a general theoretical framework for behaviour. Instead, there are disparate theories in different domains, built from inductive generalizations. In this thesis, we focus on decision-making under risk and on inter-temporal decisions (delayed rewards). We provide a background of these two areas, as well as briefly summarize some other important areas like social preferences, strategic interactions, and neuroeconomics for the sake of completeness.

Decision Making under Risk and Uncertainty A number of anomalies related to the axioms and assumptions of expected utility theory (EUT) have been observed in both lab data, and more prominently in field data (Camerer (2004)). Two examples predate prospect theory (PT), and were introduced by Allais (1953). In the *Allais Paradox*, objective probabilities of the states of the world are known, i.e. risky situations. The *Ellsberg paradox* involves ambiguity or *uncertainty* where objective probabilities cannot be calculated.

Prospect Theory (PT) was developed by Kahneman and Tversky (1979), and then extended to cumulative prospect theory in Tversky and Kahneman (1992). PT models choice as a two-phase process: editing and then evaluation. The editing phase consists of a preliminary analysis of the prospects, which yields a simpler representation and facilitates the evaluation process for the prospects. Editing involves heuristic rules and operations like coding, combination, segregation, cancellation, simplification and the detection of dominance[2].

After the editing phase, the decision maker evaluates each of the edited prospects, and chooses the one with the highest value. The value V of a prospect $(x, p; y, q)$ (first outcome is x with probability p, second outcome is y with probability q) is: $V(x, p; y, q) = \pi(p)v(x) + \pi(q)v(y)$. Where $v(0) = 0, \pi(0) = 0, \pi(1) = 1$.

The first scale assigns to each outcome x a subjective value $v(x)$[3]. The valuation function $v(x)$ depends on four parameters: the reference point (r), coefficient for diminishing marginal sensitivity to gains (α) and to losses (β), and a coefficient for loss-aversion (λ). The second scale $\pi(p)$, for decision-weighting involves an inverted-S shaped curve[4].

In PT, outcomes are defined relative to a *reference point* (i.e. a zero point of the value scale); the variable v

[2]

[3] $v(x) = (x - r)^{\alpha}$, if x $\geq r$, and $v(x) = (x - r)^{\beta}$, if x ¡ r.

[4] $\pi(p) = \dfrac{p^{\gamma}}{(p^{\gamma} + (1 - p)^{\gamma})^{1/\gamma}}$

measures deviations from the reference point, i.e. gains and losses. Kahneman and Tversky (1979) held this assumption to be compatible with the basic principles in psychophysics of perception and judgement.

Kahneman and Tversky (1979) note, "A salient characteristic of attitudes to changes in welfare is that losses loom larger than gains.". Generally, this means, $v(x) < v(-x)$, where $x > 0$. In Tversky and Kahneman (1992), the median value of the loss-aversion λ for student subjects was 2.25. This phenomenon of loss-aversion is one of the most robust findings from behavioural economics, and has been documented in the lab, in the field, and empirically, Camerer (2005). Loss aversion also makes sense from an evolutionary psychological point of view. Pinker (1999) proposed that whilst gains can improve our prospects of survival and reproduction, significant losses can take us completely 'out of the game'.

Loss aversion can be applied to explain *endowment effects*: the difference between prices potential buyers are willing to pay (WTP) for goods and prices potential sellers are willing to accept (WTA) for the same goods, as studied in lab experiments by Kahneman *et al.* (1990). *Disposition effect* is a phenomenon in finance, where investors tend to hold on to stocks that have lost value (relative to purchase price) for too long, while eagerly selling stocks that have risen in price. According to EUT, people should buy or sell based on expectations of future prices, not the past prices. A study by Odean (1998) indicated that unsold losers only yielded a return of 5% in the following year, compared with a return of 11.6% for winners that were sold later. Genesove and Mayer (2001) reported a similar disposition effect in the housing market. Owners appear to be unwilling to sell their properties for less than what they paid for them, and tend to hold on for too long before selling when the market goes into a downturn.

Kahneman and Tversky (1979) proposed a utility function that featured diminishing marginal sensitivity in the domain of gains and losses, a well-known feature of the standard model. In PT, the value function for changes of wealth is normally concave above the reference point and convex below it. This type of function implies risk-aversion in the domain of gains and risk-seeking in the domain of losses.

In prospect theory, the value of each outcome is multiplied by a decision weight. However, decision weights do not obey the probability axioms and should not be interpreted as degrees of belief. The decision weights, π, of Kahneman and Tversky (1979) had the following characteristics: π is increasing in p with $\pi(0) = 0$ and $\pi(1) = 1$. Furthermore, π violates the normal probability axioms in EUT: it is subadditive ($\pi(rp) > r\pi(p)$ for $0 < r < 1$), subcertain ($\pi(p) + \pi(1-p) < 1$) and subproportional, i.e. violates the independence or substitution axiom of EUT.

Certain empirical anomalies were later observed and the revised cumulative prospective theory was developed in Tversky and Kahneman (1992). The essential difference is that the principle of diminishing returns is now applied to weighing functions as well as the utility function. Diminishing sensitivity gives rise

to a weighing function that is concave near 0 and convex near 1. Furthermore, decision weights are assigned cumulatively, starting with the largest gain and working downwards. Although PT is less parsimonious than EUT (and its conventional extensions), Camerer (2004) observed in a review of PT with other theories, that cumulative PT can not only explain the same observations that EUT can, but also the various anomalies that cannot be explained by EUT.

Time Preference Intertemporal decision-making relates to situations where people are faced with choosing whether to undertake an activity that involves a cost sooner and a benefit later, e.g. investment goods like education, or leisure goods like watching TV or eating junk food. The long-term benefits in these situations normally exceed the short-term ones, the basic trade-off is between 'smaller-sooner' (SS) versus 'larger-later' (LL). Experimental studies have shown the phenomenon of time-inconsistent preferences. Ainslie and Haslam (1992) reported that a majority of subjects would prefer $100 certified check available immediately over $200 certified check that cannot be cashed before 2 years; yet the same people do not prefer $100 in 6 years to a $200 in 8 years. This implies that the discounting rate used by the subjects is greater over the short time horizon than over the long time horizon. Declining discount rates form the basis of hyperbolic discounting; people tend to be more impatient in the short run, using a higher discount rate, and become more patient over longer periods of time. This phenomenon is referred to as *present bias*.

The original hyperbolic discount function based on experimental studies with animals, was introduced by Chung and Herrnstein (1967), and took the form $D(t) = 1/t$. Phelps and Pollak (1968) used a modified version to study intergenerational altruism, called the *quasi-hyperbolic* function, where $D(t) = \beta\delta^t$ if $t > 0$ and $D(t) = 1$ if $t = 0$. This function is also called the (β, δ) model. The pure, generalized hyperbolic function was originally introduced by Harvey (1986), and further developed by Loewenstein and Prelec (1992). This is a continuous function of the form $D(t) = (1 + \alpha t)^{\beta/\alpha}$. The α coefficient determines how much the function departs from constant discounting, and in the limiting case of $\alpha \to 0$ is the exponential discount function. The primary implication of hyperbolic discounting is that time preferences will be inconsistent.

Several other models of time preference have been proposed in literature. Ho *et al.* (2006) illustrate the difference between exponential discounters,"naive" and "sophisticated" hyperbolic discounters. Naive agents think that they will use a constant discount rate in the future, but actually discount hyperbolically. Sophisticated agents accurately predict how their preferences will change over time and pre-committ to certain courses of action that prevent them from yielding to a preference reversal later[5]. The class of models discussed so far

[5]Ho *et al.* (2006) give a hypothetical example of the situation of buying and consuming potato chips. Exponential discounters may buy a large bag to benefit from quantity discount but only consume a single serving to avoid worst health outcome. Naive hyperbolic discounters buy a large bag believing they will only consume a single serving. However in the second period they discount hyperbolically and end up consuming in excess. Sophisticated hyperbolic discounters may choose the small bag as a commitment device, knowing that they would be unable to resist consuming extra if they bought the large bag.

use a subjective perception of money (as a function of time), while the perception of time is linear. Another class of models, with roots in biology and psychology, models the subjective perception of time which leads to inconsistencies in inter-temporal choices, e.g. Ebert and Prelec (2007), Ray and Bossaerts (2011). We present a hyperbolic discounting theory where the subjective perception of time is not linear in Chapter 3.

Several empirical studies provide evidence for hyperbolic discounting and measure the parameter values for (β, δ) model. While these studies were designed to test nonconstant hyperbolic discounting versus constant exponential discounting, they did not test hyperbolic discounting against other behavioural theories. DellaVigna and Malmendier (2006) studied health club memberships an found that a model with $\beta = 0.70$ and $\delta = 0.9995$ best fit their data. Skiba and Tobacman (2008) allowed for partial naivete in modeling decisions of borrowers of pay-day loans and found they had very high discount rates (estimated $\beta = 0.53$, $b = 0.90$ and $\delta = 0.45$).

Although this thesis looks at individual decision making, in particular risky choice and prospect theory, and time preferences, some major developments in behavioural economics have also come from behavioural game theory, social preference and neuroeconomics. We only provide a cursory overview here, which is necessary for the sake of a complete introduction.

Behavioral Game theory There are four main elements of behavioural game theory (BGT) corresponding to assumptions in the standard game-theoretic model (SGT) (Camerer (2003)): (i) Representation: Often players perceive a game incorrectly or incompletely. (ii) In SGT, player's beliefs about the game is correct since they have unbounded rationality. BGT takes bounded rationality into account, either through limits on strategic thinking (as in cognitive hierarchy theory, Camerer *et al.* (2004)), or by assuming players make stochastic mistake because of 'noise'. (iii) Learning (which is ignored in SGT) is relevant in repeated games where players can learn from their own and other's payoffs, other player's strategies, and also what other players are likely to do, Camerer (2003). (iv) Players have preferences regarding not only their payoffs but also those of others, and the distribution of those payoffs.

Social Preference In contrast to the standard model which assumes that economic agents are motivated purely by self-interest, behavioural economics points to a large amount of evidence that people have social preferences. According to Fehr and Fischbacher (2005), for social preferences, not only are payoffs relevant, but also beliefs and intentions can be valued positively or negatively. Three sylized games: ultimatum bargaining, dictator and trust games, played in both the lab and in the field, across different cultures and countries, and with very high and low stakes, have provided strong evidence of this phenomenon. Fehr and Schmidt (1999) proposed an inequity-aversion model, sometimes referred to as 'envy/guilt' model where

fairness judgements are based around some 'neutral' reference point. Social psychology research, e.g. Adams (1963), has shown that relative material payoffs affect people's wellbeing and behaviour; A number of empirical findings also support the hypothesis, e.g. Loewenstein *et al.* (1989). Reciprocity models, first introduced formally by Rabin (1993), are based on perceived intentions that determine positive or negative reciprocity.

Neuroeconomics Neuroeconomic studies rely on brain scanning and imaging techniques, including PET (positron emission tomography), fMRI (functional magnetic resonance imaging), EEG (electroencephalography) and TMS (transcranial magnetic stimulation). Other simpler biological measurements also fall under the purview of Neuroeconomics: eye-tracking and skin-conductance response. But, despite the growing evidence and contributions, Neuroeconomics has not been without its share of controversies. Gul and Pesendorfer (2008) are the most forceful critics who claim that if two different economic models make different predictions regarding decisions then they can be evaluated by examining standard choice data; if the two models make the same predictions, then economists will not be interested in distinguishing between them. To a scientist, this argument is unsatisfying - the goal of science is to create progressively realistic models that can explain different types of data, which only improves its out-of-sample predictive powers; there's no need to force restrictions. Bernheim (2008) criticizes neuroeconomics since it is a reductionist approach. Glimcher *et al.* (2009) argues that the neuroscientific findings act as constraints on theories, that can be viewed as compatible with human neuroarchitecture[6].

De Martino *et al.* (2006a) examined the differential responses of subjects to losses and gains in gambling tasks and found support for the existence of framing effects and reference dependence, with differences in orbitofrontal cortex (OFC) and amygdala modulating behaviour. For a more causal relationship, De Martino *et al.* (2010) found that patients with amygdala lesions did not exhibit loss aversion. fMRI studies have also been used for comparing different behavioural models, e.g. exploration and exploitation in decision making under uncertainty (Daw *et al.* (2006)), strategic interactions and social preferences (Xiang *et al.* (2012)).

A common criticism of behavioural economics is that there are too many theories, each explaining different phenomenon under different contexts. Our aim in the thesis is to compare these theories with different types of data. Alongside field studies and lab experiments, neuroeconomics - combining neuroscience and neurobiology - can also provide evidence in support of a theory over others.

[6]Bernheim (2008) argues against a reductionist approach by saying 'Do we really believe that good economics requires a command of string theory?'. This isn't a valid argument since physical laws constrain or allow the space of possible human activity, and thus constrains economics. For example, the speed of light naturally imposes limits on certain economic activity. Were it not so, some absurd economic realities might arise, like intergalactic arbitrage (Krugman (2010)).

1.2 Experimental Economics

Accounting for the history of experimental economics, Roth and Kagel (1995) used the example of Daniel and Nicholas Bernoulli on the St. Petersburg paradox (Bernoulli (1738)). The early periods of neoclassical economics (the orthodox approach to the subject for most of 20th century) was based on experimental research. In launching the "marginal revolution" in economic theory, Jevons (1870) and Edgeworth (1881) based their analysis of diminishing marginal utility on psychological findings by psychophysicists, such as Fechner (1860), about the relationship between stimuli and sensations.

In a now classic experiment, Thurstone (1931), wanted to provide empirical grounding of the concept of an indifference curve in economic theory. His experiment attempted to elicit individual's indifference curves from responses to binary choice problems. Over the next three decades, empirical experiments to test whether the preferences postulated in theory could be elicited from actual choice data, was scant (for e.g. Allais (1953), Davidson and Marschak (1959)). Chamberlin (1948) investigated price-determination in an experimental market, motivated by his awareness that price theory provided no real explanation of how equilibrium was achieved in real markets. His results showed that equilibrium would not be reached under conditions typical of real-world markets.

Smith (1962) reported a series of experiments that refined and extended Chamberlin (1948)'s first market experiments, to test hypotheses derived from competitive equilibrium price theory. The strategy was to create an experimental market setting with demand and supply conditions that are known to the experimenter, and stable over a sequence of market periods. Smith saw stable demand and supply as atypical of real markets, but built into the laboratory experiments to test a particular economic hypothesis: that behaviour will converge to equilibrium in a stable environment. The experiments involved sequences of market periods where buyers and sellers could make bilateral contracts via an oral double-auction institution. Smith observed a remarkable degree of convergence of prices, across periods, to the equilibrium prediction.

Smith *et al.* (1988) reported several experiments investigating experimental markets in which participants could trade "cash" or "assets" over a series of (usually fifteen) market periods. The markets were organized as computerized double auctions. Variations across the expeirments explored various issues, including the extent to which "experience" level of traders affected conformity of market behaviour with predictions. The authors reported that it was common for their markets to exhibit "bubbles", i.e. sustained periods where assets traded at prices significantly above their expected returns. Such bubbles were typically followed by crashes, with prices and volume of trades collapsing near the final period. Trader's experience also had some tendency to attenuate bubble phenomena.

The 2002 Nobel prize awarded to Daniel Kahneman and Vernon Smith led to a popularization of ex-

perimental economics. It is also suggested that the emerging widespread use of experimental methods in economics has become possible through recent developements in information technology. Some experimental designs routinely used nowadays would be infeasible a few decades ago. Generic software, such as z-Tree package by Fischbacher (2007) makes it easier to design and run experiments. Adaptive experiments will also make experimental economics more accessible to a wider group of researchers by decreasing the time and costs associated with experiments. El-Gamal and Palfrey (1993) introduced optimal endogeneous experimental designs into experimental economics in the 1990s, which did not gain traction at the time. The time is now ripe for a resurgence in adaptive experiments due to order of magnitude increase in computational power (Camerer (2003)), and recent algorithmic advancements[7].

External Validity Conditions in experiments are typically markedly different from those in naturally occurring situations. This might give rise to uncertainty and even skepticism about how, or whether, the results will generalize. Doubts about external validity are usually expressed as respects in which laboratory experiments are "artificial". Smith (1962) illustrates one of the key features of laboratory testing: the possibility to control the environment to identify and manipulate the central variables of a theory. Smith (1982) outlined a set of "precepts" that, if satisfied, would be sufficient to guarantee that the experimenter has induced the intended set of preferences[8].

The basic precepts are *nonsatiation, saliency, dominance* and *privacy*. Nonsatiation requires that subjects prefer more than less of the reward and do not become satiated by the range of possible rewards in the experimental setting. Saliency is the requirement that rewards are appropriately tied to decisions in the experiment; Specifically, rewards should be monotonically increasing in the objective the experimenter is trying to induce. Dominance requires that other costs (e.g. transaction costs) do not override the motivation provided by the reward in the experiment. Privacy requires that a subject knows only his or her own reward, in order to rule out interpersonal considerations such as envy or altruism[9].

Our laboratory experiments satisfy all the above precepts. Subjects are paid with money (nonsatiation) based on their decisions in the task (saliency). We pay particular attention to removing transaction costs, by paying a show-up fee to compensate subjects for their time, and participants make decisions without any knowledge of other's decisions.

[7] Camerer notes enthusiastically in pg 42 of Camerer (2003), "Furthermore, because of increases in computing power, for the first time in human history we can alter the experimental design in real time - while subjects are waiting, for seconds rather than days - to optimize the amount of information collected in an experiment."

[8] Experiments satisfying this *induced-value methology* has become a standard technique to implement particular sets of preferences in market experiments.

[9] In laboratory experiments investigating social preference, the payoffs are explicitly tied, and the participants remain anonymous.

Incentives While the standard in experimental psychology might be to rely on subject's intrinsic motivation to perform a task, the norm in experimental economics is to compensate subjects for their time (opportunity cost) as well as compensating according to performance. In our experiments, subjects' earnings corresponds entirely to their performance in the task based on choices made. Additionally the earnings from the experimental are expected to be higher than the typical wage rates earned by the subjects.

An important departure between experiments in economics and other fields that study decision-making, such as experimental psychology, is the attitude towards *deception*. Proponents of the use of deception argue that deceiving subjects could be a convenient way of achieving experimental objectives (see Bonetti (1998)), such as making subjects believe that they are in situations that would be difficult to engineer by other means.

Most economists, however, are in favour for prohibiting deception in experiments (Hertwig and Ortmann (2001)). The case largely rests on the argument that trust between subjects and experimenters is a public good, and even limited use of deception creates a negative externality.

Furthermore, there is reason to believe that the internal validity of experimental research depends on subjects trusting the experimenters enough to believe what the experimental instructions say about the decision environment. The position in this thesis is to avoid deception altogether in our experiments. Subjects are clearly informed of the structure and variables in the experiment[10].

1.2.1 Errors in Lab Experiments

Several frameworks for choice have emerged from behavioural economics, as noted previously, and there haven't been enough tests to distinguish between models. This is primarily due to economic constraints of time and money. Models, such as prospect theory, have multiple parameters and would require a very large number of choices to distinguish by standard experimental design and statistical techniques. This makes it infeasible to test in practice since subjects' time is limited[11]. For this reason, researchers in behavioural economics have been enthusiastic about adaptive experimental designs[12], that can test multiple theories with few choices.

There is a hypothetical possibility that subjects, when presented with adaptive tests might engage in strategic manipulation which would bias the results, or in some way distort their actual behaviour. This criticism could pose a hindrance in the adoption of such methods in economics, so we investigate this issue

[10] Since economic theories tell us almost nothing about how big incentives need to be to motivate participants, experiments must make assumptions about the level of rewards that are necessary. We follow the suggestion of Smith: "payoff levels that are judged to be high for the subject population", Smith (1982).

[11] It is feasible to run multiple sessions, but subjects' attention deterioriates while performing repeated tasks. And also subjects need to be compensated for their time, and this imposes financial limitations on researchers.

[12] In Camerer (2003), Camerer mentions: "In many experiments, the experimenter has one more hypotheses that she can put prior probabilities on. A crisp prior, and specification of the hypotheses, can be used to compute the information value of different experimental design parameters. This motivates the choice of "optimally informative" design parameters... Seen this way, all previous experimental designs are heuristic approximations to endogenously optimized designs. Information-optimized designs have rarely been used.".

in depth in chapter 1. To summarize, the main arguments against subjects strategically misreporting their answers for higher rewards are: a) Subjects do not know how tests are sequentially selected b) Even if they were aware of the method used, strategizing would be computationally infeasible c) The increased earnings of a clairvoyant agent with boundless computational ability is barely signifcant, and d) We see no evidence for strategizing in our data.

Hypothesis testing methods in Econometrics pay attention to sources of error in (non-experimental) observational data. This includes structural or model specification error, measurement error and sampling error. Because experiments are carefully designed to ontrol many of the other sources of noise in non-experimental data, variability within and between actors' beliefs, preferences and judgements play a more central role in the generation of experimental data. In decision making under risk in the nonexperimental world, people operate on their own evaluations of consequences and their own subjective estimates of risks, which vary in ways that are hard to observe. Individual decision-making experiments, on the other hand, involve a small number of clearly specified payoffs and use probabilities corresponding to some explicit random mechanism. Thus, the principal component of variation in the data is likely to derive from participants themselves[13]. Modeling the stochastic component, and using appropriate tests, requires us to model the nature of noise and error in human judgement and decision making. Three different "error stories" have been popular in economics, namely the *tremble* model, *fechner* model and the *random preference* model.

The tremble model captures the simplest source of error, where people have "true" preferences, but occassionally make mistakes in reporting them due to momentary inattention, or in short a "trembling hand". Formally, an individual truly prefers one option, but there is some probability ϵ that she will choose the less-preferred option while translating preference into choice. Harless and Camerer (1994) use a trembling hand model to facilitate a meta-analysis of studies of violations of independence. The Fechner model, originating in psychophysics ascribes the error to imperfections in human judgement. The true net advantage of an option f over option g is $V(f, g)$, bu the perceived advantage at the moment of choice is given by $V(f, g) + \epsilon$, where ϵ is a stochastic error term. The variance of ϵ is often an increasing function of the magnitude of the stimuli. Instead of supposing a single true preference function to which noise is added, the random preference approach (Loomes and Sugden (1995)) supposes that an individual's preferences consist of a *set* of such functions. Intuitively, an individual's perceptions and judgements are liable to vary to some extent from moment to moment along some spectrum of "states of mind".

Our BROAD method pays particular attention to "noisy responses" or within-subject error. Unlike other objective criteria used in literature, our criteria is robust to some level of within-subject error, i.e. the

[13]Loomes and Sugden Loomes and Sugden (1998) did a study invovling 92 respondents each making 40 choices on two occassions, where the same pair of binary choices between two lotteries was presented for a second time. Out of a total of 3680 comparisons, there were 676 (18.4%) cases where the choice on the second occassion was different from the first.

conclusions reached remain unaffected by some response error. Furthermore, we provide the first theoretical guarantees of (near-)optimality for sequential experimental design in the presence of noise.

1.2.2 Field Experiments

The gold standard in field settings exploit natural experiments[14] in which there is a true random assignment, or there is a good instrumental variable (Angrist and Krueger (2001)) to achieve control. Wonderful instruments are often created by unusual occurences in policy changes or historical accidents. In Chapter 3 we exploit natural variation in prices, in particular a retailer's goods being on or off-promotion, to test unique predictions of prospect theory.

Levitt and List (2007) criticize the inability of lab experiments to generalize and argue that field experiments are more reliable since they are superior in terms of external validity[15]. Camerer (2011) provides a critical reply to Levitt and List (2007): two criticisms i) presence of moral and ethical considerations and ii) nature and extent of scrutiny of one's actions by others are equally valid for lab and field experiments. We see the third criticism iii) context in which decision is embedded, as a feature of lab experiments rather than a negative issue - understanding how context affects decisions is essential in adding psychological realism to models, and lab experiments can control the context effectively. Two other criticisms iv) self-selection of individuals making the decision, and v) the stakes of the game, are valid for lab experiments, and can be corrected by experimenters taking the necessary precautions. Camerer (2011) further notes that lab-field generalizability is the norm rather than the exception[16].

We maintain the *scientific view* in this thesis, that all empirical studies contribute evidence about the general way in which agents characteristics, incentives, rules, information, endowments and payoff structure influence economic behavior. In this thesis, we find evidence for Prospect theory, especially reference dependence and loss aversion in both our (laboratory) experiments and (natural) field settings, which contributes to the growing evidence in literature. Phenomena that are evidenced in both lab and in the field have greater empirical validity, and prospect theory (in particular loss aversion) has been found in both settings.

[14] Friedman (1953, pg 10) noted about natural experiments: "Unfortunately, we can seldom test particular predictions in the social sciences by experiments explicitly designed to eliminate what are judged to be the most important disturbing influences. Generally, we must rely on evidence cast up the "experiments" that happen to occur.

[15] Levitt and List (2007) expressed their criticm in a rather provocative way: "Yet unless considerable changes are made in the manner in which we conduct lab experiments, our model highlights that the relevant factors will rarely converge across the lab and many field settings".

[16] Camerer (2011) finds that 20 studies confirm lab-field generalization and only 2 that disqualify.

1.3 Methods

We have outlined the research questions and methodologies from behavioural and experimental economics above. Here we give a brief background of the methods (primarily from Computer Science and Econometrics) that are developed and utilized in this thesis. First we introduce the mathematical concept of submodularity, and its properties, which is used extensively in chapters 2 and 3.

1.3.1 Submodularity and Greedy Algorithms

Submodularity is a property of set functions (function $f : 2^X \rightarrow R$) that assigns a value $f(S)$ to each subset $S \subset X$. V is a finite set, which could refer to locations where observations are taken, for example, in a sensor network. $f(S)$ is the utility obtained from taking observations at S. Submodularity can be defined in terms of discrete derivate:

For a set function $f : 2^X \rightarrow R$, $S \subset X$, and $x \in X$, let $\Delta_f(x|S) \equiv f(Sx) - f(S)$ be the discrete derivate of f at S with respect to x.

Equivalently submodularity can be defined in terms of marginal gain:

A function $f : 2^X \rightarrow R$ is submodular if for every $A \subset B \subset X$ and $x \in X$ B it holds that $\Delta(x|A) \geq \Delta(x|B)$. Or equivalently $f(AB) + f(AB) \leq f(A) + f(B)$.

Intuitively, submodular set functions exhibit a natural diminishing returns property. An important subclass of functions exhibit monotonicity:

A function $f : 2^X \rightarrow R$ is monotone iff all its discrete derivatives are nonnegative. The subclass of monotone submodular are characterized by the fact that for all $A \subset B \subset X$ and $x \in X$ it holds that $\Delta(x|A) \geq \Delta(x|B)$. An example of a monotone submodular function is the Shannon entropy function $H(X_S) = -\sum_{x_S} P(x_S) log_2 P(x_S)$, where $P(X)$ is a joint probability distribution over discrete random vector X, as proved by Fujishige (1978). The mutual information: $f(S) \equiv I(Y; X_S) = H(Y) - H(Y|X_S)$, for a joint probability distribution $P(X, Y)$ over two dependent random vectors, X and Y quantifies the expected of uncertainty about Y after revealing X_S. In general, f is not submodular. However, if the variables X and Y are conditionally independent given Y (a Naive Bayes model), then f is monotone submodular (**?**)).

There are many natural connections between submodular, concave and convex functions (explored by Lovász (1983)). For a function $g : N \rightarrow R$, the set function $f(S) = g(|S|)$ is submodular if and only if g is concave. Similar to convex functions, which can be minimized efficiently, submodular minimization is possible in (strongly) polynomial time (Schrijver (2003)). Furthermore, submodularity is closed under nonnegative linear combinations and residuals.

We are interested in maximization of submodular functions, i.e. $max_{S \subseteq V} f(S)$ subject to some constraints

on S. For example, S can have cardinality constraints (e.g. constrained budget) where $|S| \leq k$ for some k. This problem is NP-hard for several classes of submodular functions, including mutual information (**?**)). Hence efficient algorithms are required with theoretical approximation guarantees.

A simple approach of solving submodular maximization with cardinality constraints is the greedy algorithm. We start with the empty set S_0, and in iteration i, add the element maximizing the marginal gain $\Delta(x|S_{i-1})$:

$$S_i = S_{i-1} \cup \{\arg\max_x \Delta(x|S_{i-1})\}$$

. A famous result by Nemhauser *et al.* (1978), restated here, proves that the greedy algorithm provides a good approximation to the optimal solution of the NP-Hard optimization problem. In fact, for several classes of submodular functions, this is the best result that can be achieved with any efficient algorithm.

Fix a nonnegative monotone submodular function $f : 2^X \to R_+$ and let $\{S_i\}_{i\geq 0}$ be selected by a greedy algorithm, that adds the element with maximum marginal gain at every iteration. Then for all positive integers k and l,

$$f(S_l) \geq (1 - e^{-l/k}) \max_{S:|S|\leq k} f(S)$$

In particular, for $l = k$, $f(S_k) \geq (1 - 1/e) \max_{|S|\leq k} f(S)$.

The min-cost coverage problem involves searching for minimum cost sets that achieve a given amount q of submodular value, i.e. $S* = \arg\min_S |S| s.t. f(S) \geq q$, for some quota $0 \leq q \leq f(V)$ of value. Wolsey (1982) proved the following result about the greedy algorithm:

Suppose $f : 2^X \to N$ is monotone submodular and integer-valued, and let $0 \leq q \leq f(X)$. Let $S_0, S_1, \cdots$ be the sequence of sets picked by the greedy algorithm, and let l be the smallest index such that $f(S_l) \geq q$. Then

$$l \leq (1 + \ln \max_{x\in X} f(\{x\}))OPT$$

where $OPT = \min_S |S| s.t. f(S) \geq q$.

Another benefit of submodular functions is that the greedy algorithm can be sped up, sometimes with orders of magnitude gains in speed, by using lazy evaluations (discussed in Chapter 2).

In several applications, such as sequential experimental design or active learning, we wish to adaptively select a set, make observations, and use the feedback after selecting any particular element. In such adaptive optimization problems we optimize over policies (equivalently strategies) i.e. functions from the information we have obtained to the next action.

We model states (of the world) as a random variable H (instances are denoted h). H could refer to the set of all hypothesis we wish to test. We presume a Bayesian model with a prior probability $P(H)$ over H. We

suppose there is a set of actions X (for example, tests we can run) we can perform and a set of outcomes or observations O. The state h is a function from actions to outcomes of those actions, i.e. $h : X \to O$ and $h(x)$ is the outcome of performing action x. The domain of ψ, denoted $\mathrm{dom}(\psi)$, is the set of actions performed up to that point. We call h a *realization* and ψ a partial realization. We assume that there is an objective function $f : 2^X \times O^X \to R_+$ indicating the reward $f(A, h)$ obtained from actions A under realization of the world state h. A policy π can then be a mapping from partial realizations ψ to the actions, so that $\pi(\psi)$ is the action taken by π upon observing ψ. Finally, define $V(\pi, h)$ to be the set of actions by π under realization h. This leads to two optimization problems:

The expected reward of a policy π is

$$f_{avg}(\pi) \equiv E[f(V(\pi, H), H)] = \sum_h P(h) f(V(\pi, h), h)$$

. The goal of Adaptive Stochastic Maximization problem is to find a policy $\pi*$ such that

$$\pi* \in \arg\max_\pi f_{avg}(\pi) s.t. |V(\pi, h)| \leq k \forall h$$

, where k is a budget over the number of actions (e.g. the number of tests we can perform).

Alternatively, we can define the average cost $c_{avg}(\pi)$ of a policy as the expected number of actions it makes, so that $c_{avg}(\pi) \equiv E[|V(\pi, H)|]$. The goal of the Adaptive Stochastic Minimum Cost problem is to find

$$\pi* \in \arg\min_\pi c_{avg}(\pi) s.t. f(V(\pi, h), h) \geq q \forall h$$

, i.e. the policy that minimizes the expected number of items picked such that under all possible realizations, at least reward q is achieved.

Both problems are notoriously intractable: even approximating a solution (to a factor of $O(|V|^{1-\epsilon})$) is intractable. However, if f satisfies certain generalizations of monotonicity and submodularity, defined below, the performance bounds can be generalized.

Given a partial realization h and an item x, the *conditional expected marginal benefit* of x conditioned on having observed ψ, is

$$\Delta(x|\psi) \equiv E[f(\mathrm{dom}(\psi)\{x\}, H) - f(\mathrm{dom}(h), H)|\psi]$$

, where the expectation is taken w.r.t $P[h|\psi]$.

A function $f : 2^X \times O^X \to R_+$ is adaptive monotone w.r.t. distribution $P[h]$ if the conditional expected marginal benefit of any item is nonnegative, i.e. for all ψ with $P[\psi] > 0$ and all $x \in X$ we have $\Delta(x|\psi) \geq 0$.

A function $f : 2^X \times O^X \rightarrow R_+$ is adaptive submodular with respect to distribution $P[h]$ if the conditional expected marginal benefit of any fixed number of items does not increase as more items are selected and their states are observed. Formally, f is adaptive submodular w.r.t. $P[h]$ if for all ψ and ψ' such that ψ is a subrealization of ψ' (i.e. $\psi \subset \psi'$), and for all $x \in X \ \mathrm{dom}(\psi')$, we have $\Delta(x|\psi) \geq \Delta(x|\psi')$. Adaptive submodularity reduces to submodularity in the case when the world state realization H is deterministic. Analogously, the *adaptive greedy algorithm* iteratively selects the action maximizing the conditional expected marginal benefit, conditioned on the outcomes of all its previous actions[17].

For adaptive stochastic maximization, the algorithm terminates after selecting k actions. For adaptive stochastic minimum cost cover, it stops after achieving the quota q of value. A key result, proved by Golovin and Krause (2010a), restated here is used to prove the results of adaptive submodular optimization with noisy observations in Chapter 2.

Let π_l^{greedy} be the greedy policy, implemented by the adaptive greedy algorithm, run for l iterations (to select l actions), and let π_k* be any policy selecting at most k actions for any realization h. Then

$$f_{avg}(\pi_l^{greedy}) \geq (1 - e^{-l/k}) f_{avg}(\pi_k*)$$

where $f_{avg}(\pi) \equiv E[f(V(\pi, H), H]$, is the expected reward of π.

Another import application for submodularity is in online learning. In the above we have assumed that the utility function f is known. In some applications the objective function may not be known in advance. If $f_1, \cdots, f_T$ is drawn from some distribution and we perform the same task repeatedly, we can learn to perform well on average over time. This is the premise behind *no-regret* algorithms that are used widely in machine learning. This approach is not used in this thesis; refer to Krause and Golovin (2013) for a review. Notably, Srinivas *et al.* (2012) proved no-regret bounds for online experimental design under certain constraints (values of f are sampled from a known Gaussian Process).

1.3.2 Discrete Choice models

Discrete choice models describe the decision makers' choices among alternatives such as competing products or actions. In the framework, the set of alternatives, called the *choice set* must be *mutually exclusive* (choosing one alternative implies not choosing any of the others), *exhaustive*, and countably finite.

The discrete choice models can be derived under the assumption of utility-maximization by the decision maker. Thurstone (1931) introduced the law of comparative judgement, from the point of view of psychophysics

[17] At each iteration of the adaptive greedy algorithm: (i) Select $x* \in \arg\max_x \Delta(x|\psi)$; (ii) Observe $h(x*)$; (iii) Set $\psi \leftarrow \psi\{(x*, h(x*))\}$

now called the binomial probit model. Marschak (1950) later interpreted the stimuli as utility and provided a derivation from utility maximization; models derived this way are called *Random Utility Maximization* (RUM) models. Although it is consistent with utility maximzation, this does not preclude the model from being consistent with other forms of behaviour.

In RUMs, the utility the decision maker n obtains from alternative j is $U_{nj}, j = 1, \cdots, J$. The decision maker chooses the alternative with greatest utility, i.e. $\arg\max_i U_{ni}$. The researcher however does not observe the utilities, only attributes of the alternatives $x_{nj} \forall j$, and some attributes of the decision maker, s_n. The *representative utility* constructed by the researcher, depending on unknown parameters, is $V_{nj} = V(x_{nj}, s_n) \forall j$. Utility is decomposed as $U_{nj} = V_{nj} + \epsilon_{nj}$, where ϵ_{nj} are random structural errors - unknown variables not captured by the model. The joint density of the random vector $\epsilon'_n =< \epsilon_{n1}, \cdots, \epsilon_{nJ} >$ is denoted $f(\epsilon_n)$. The probability that the decision maker n chooses i is:

$$P_{ni} = P(U_{ni} > U_{nj} \forall j \neq i) = \int_{\epsilon} I(\epsilon_{nj} - \epsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i) f(\epsilon) d\epsilon_n \tag{1.3.1}$$

where I is an indicator function, and the multidimensional integral is over th density of the unobserved portion of utility, $f(\epsilon_n)$. Different discrete choice models result from different different assumptions over $f(\epsilon_n)$. Train (2009) provides a comprehensive overview of the models and underlying assumptions.

The simplest and most widely used discrete choice model is the logit, which has a closed form solution. If the representative utilities are linear in parameters, i.e. $V_{nj} = \beta' x_{nj}$, then the logit probabilities become:

$$P_{ni} = \frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}} \tag{1.3.2}$$

The log-likelihood function of these choice probabilities is globally concave in β which can be maximized with simpler numerical procedures like the Newton-Rhapson method.

In terms of identification, only differences in utility matter, and the overall scale of the utility is irrelevant. For any two alternatives i and k, the ratio of logit probabilities is $P_{ni}/P_{nk} = e^{V_{ni} - V_{nk}}$. The ratio does not depend on any other alternatives than i and k. Thus the logit model exhibits *Independence from Irrelevant Alternatives* or IIA. There are situations where the unobserved portions of utility are correlated. The probit model assumes jointly normal unobserved utility components, and the correlations between the errors can be specified in the covariance matrix.

Mixed multinomial logit (MMNL) can approximate, to an arbitrary degree, any regular RUM-consistent

choice probabilities, under very mild restrictions on the class of choice sets (McFadden and Train (2000)).

$$P_c(i) = \int_0^1 \cdots \int_0^1 \frac{e^{Z_i \alpha(\epsilon)}}{\sum_{j \in C} e^{Z_j \alpha(\epsilon)}} d\epsilon \tag{1.3.3}$$

where $\alpha(\epsilon)$ is a vector of polynomial functions of the uniform random vector ϵ and Z_j are vectors of polynomial functions of observed characteristics of the consumer and observed attributes of alternative j. This is a *latent variable* model and has ties to computational neuroscience and machine learning. This class of latent variable models is a *single hidden-layer feedforward neural network* with MNL activation functions. The asympototic approximation theory developed for neural networks can be used to prove convergence rates and stopping rules (Cheng and Titterington (1994)). Furthermore Bayesian methods, developed for neural networks (Neal (1995)), can be utilized to learn the parameters and do inference on MMNL models.

As noted by McFadden (2001) in his Nobel prize lecture, the random utility maximization (RUM) theory of decision-making is flexible enough to model psychological factors in consumer behaviour. It is thus a useful extensible framework for incorporating non-choice measurements and elements of behavioural economics.

1.4 Findings and Contributions

Here we summarize the central contributions of this thesis:

- In chapter 2, we introduce the methodology of Bayesian Rapid Optimal Adaptive Designs (BROAD). BROAD sequentially chooses the "most informative" test at every step, and updates the beliefs over hypotheses based on (possibly noisy) responses, which informs the selection of the next test to run. We prove that the Equivalent Class Edge Cutting, EC^2, criteria used by BROAD satisfies the mathematical property of adaptive submodularity. This allows us to prove theoretical guarantess against the intractable Bayes-optimal solution, even in the presence of noisy responses. In simulated ground truth experiments, BROAD uses fewer tests than Information Gain and Generalized Binary Search. Furthermore, we use the property of adaptive submodularity to implement an accelerated greedy version of BROAD which leads to orders of magnitude speedups. BROAD thus enables us to conduct real-time laboratory experiments, to compare theories with multiple parameters requires complex experimental designs with very large possible choice sets, which are infeasbile with classical experimental design.

- In chapter 2, we apply BROAD to compare the main classes of theories for decision-making under risk: expected value, prospect theory, constant relative risk aversion (CRRA) and moments' models. 57 subjects from Caltech and UCLA were presented with choices between two lotteries, with realizable

losses. We find limited evidence in favour of CRRA and moments models from aggregate posterior model probabilities, and most subjects are classified by prospect theory, followed by expected value. Strategic manipulation, whereby subjects hide their true preferences to create a more favourable sequence of tests, is a possibility in adaptive experimental design. We give reasons why strategic manipulation is infeasible in practice, and use an experimental treatment that encourages manipulation. We find no signatures of strategizing in the data.

- In chapter 3, we compare the main theories of time preference or discounting models: exponential, hyperbolic, "present bias" i.e. quasi-hyperbolic (α, β) and fixed cost model, and generalized-hyperbolic discounting. 40 subjects from UCLA were given choices between a smaller but more immediate payoff versus a larger but later payoff. We found very limited evidence for present bias models and hyperbolic discounting, and most subjects were classified as generalized hyperbolic discounting types, followed by exponential discounting.

- The time preference models compared in chapter 3 assumes a linear mapping between the subjective perception of time and the calendar time. We consider a psychological model where the perception of time is subjective, and prove that when the biological (subjective) time is positively dependent, it implies hyperbolic discounting and temporal choice inconsistency.

- Motivated by the evidence for prospect theory in laboratory experiments, we sought to test the predictions of behavioral theories in the "wild" in chapter 4. Loss aversion and reference dependence predicts that consumers will behave in a distinct way than the standard rational model. Specifically, loss aversion distinctly predicts that when an item is being offered at a discount, the demand for it will be greater than that explained by price elasticity. Even more importantly, when the item is no longer discounted, demand for its close substitute would rise in excess. In contrast, the standard model predicts that consumers might "stockpile" when an item is discounted, and there will be no change (or a possible decrease) in demand for the substitute when the discounting period for an item is over. We tested this prediction using a discrete choice model with loss-averse utility function on data from a large eCommerce retailer. Not only did we identify loss aversion, but we also found that the effect decreased with consumers' experience.

Chapter 2

Bayesian Rapid Optimal Adaptive Designs

Choices used in experimental and survey-based social science are typically developed by hunches and cumulative search for informative questions that test theories. The conventional designs that have emerged are typically a fixed set of test questions .[1]

A popular example in estimating risk preference, introduced by Holt and Laury (2005) (HL, hereafter)[2] is a "choice list", in which one makes a sequence of ten choices between two gambles that each have the same set of possible payoffs throughout the sequence, but in which the distribution over payoffs varies. Specifically, each gamble has two payoffs, high and low, and the probability of the high payoff is $i * 0.1$ for both gambles in the i^{th} choice in the list. Using a fixed set of items permits rapid near-replications to establish rugged stylized facts across studies.

We propose an approach, in which the sequence of choices is customized for each subject rather than fixed. The subjects themselves tell us, through their answers, the "best"(most informative) question to ask them next. We illustrate the approach using an experiment on theories of risky choice: Viz., expected utility, prospect theory, and statistical moments models (mean-variance-skewness). Our approach is abbreviated as BROAD, for **B**ayesian **R**apid **O**ptimal **A**daptive **D**esign. The BROAD method is an innovation in an old, large family of adaptive methods (described in the next subsection). The big contribution is an edge-cutting measure of information value which is adaptively submodular, which therefore provably guarantees some useful theoretical and practical properties. The method was introduced by Golovin *et al.* (2010), and applied here to novel economic questions, and includes new ideas about how to detect and control strategizing by subjects.

[1] Note that "test questions", in experimental economics could include choices from budget sets Andreoni and Miller (2002); Choi *et al.* (2007); Fisman *et al.* (2007), strategies in games, auction bids, and trading strategies. Our method could also be applied to large-scale panel surveys in which branching methods are already used (as in, "If your answer to (18) is "No" please skip to question (24)").

[2] The HL paper had 1764 Google Scholar citations as of 8/27/2012.

Earlier applications of adaptive methods were made in statistics (Lindley (1956)), decision theory (Howard (1966)), computer-assisted testing (CAT) in psychometrics (e.g., Wainer (1990)) and cognitive psychology (e.g., Myung and Pitt (2009)), adaptive choice-based conjoint measurement in marketing (e.g., Abernethy *et al.* (2007)), "active learning" methods in computer science (e.g., Golovin and Krause (2010b)) and machine learning (Nowak (2009); Dasgupta (2004)), and management science (Cavagnaro *et al.* (ming); Toubia *et al.* (ming)).

Note that some early efforts to introduce optimal adaptive design in experimental economics (e.g., El-Gamal *et al.* (1993), El-Gamal and Palfrey (1993), Moffatt (2007)) did not gain traction. The time is now riper for BROAD methods because: Computing power is better than ever; a new method from computer science (called EC^2) applied here provides theoretical guarantees on efficient computability; and there are many new competing theories in behavioral economics which need to be efficiently compared.

In our application and all the others mentioned above, the goal is to sequentially select among a set of noisy, expensive observations (outcomes of experiments, medical tests to run to diagnose patients, etc.) in order to determine which hypothesis (theory, diagnosis, etc.) is most accurate. One way to formalize such active learning problems is *Bayesian experimental design* (Chaloner and Verdinelli (1995)). This approach specifies prior beliefs over a set of hypotheses, as well as probabilistic assumptions about the outcomes of test choices. The goal then is to determine the correct hypothesis while minimizing the cost of the experimentation (where cost is synonymous, for present purposes, with the number of test choice questions that are asked). Unfortunately, finding an optimal sequence of test choices is not just computationally difficult (NP-hard) but is also difficult to approximate (Chakaravarthy *et al.* (2007)). Several heuristic approaches have been proposed that perform well in some specific applications, but do not have theoretical guarantees (e.g., MacKay (1992)); that is, there are no proofs about how costly the heuristic sequence will be compared to the optimal sequence.

In the case where observations are *noise-free*[3], a simple algorithm, *generalized binary search*[4] *(GBS)*, is guaranteed to be competitive with the optimal test sequence, since the expected number of tests is a factor of $O(\log n)$ (where n is the number of hypotheses) more than that of the optimal policy (Kosaraju *et al.* (1999)), which matches lower bounds up to constant factors[5] (Chakaravarthy *et al.* (2007)).

Unfortunately, when responses are likely to be imperfect or *noisy*, as in economics experiments, the theoretical basis and performance of heuristic methods is not well understood. There are no efficient algorithms that provably achieve minimal cost, except in very restricted settings (Krause and Guestrin (2009)). [6]

[3]This case is known as the *Optimal Decision Tree* (ODT) problem.

[4]GBS selects test choices to maximize, in expectation over the test outcomes, the probability mass of eliminated hypotheses (i.e., those with zero posterior probability, computed with respect to the observed test outcomes). The method is myopic (or "greedy" in computer science terms) because it only searches ahead for one test at a time.

[5]In other words, unless P = NP, no efficient algorithm can be guaranteed to find a policy whose expected number of tests is less than C log(n) times that of the optimal policy, for some absolute constant C (see Papadimitriou (2003)).

[6]While there are some recent positive results in understanding the label complexity (bounds on the sample size required) of noisy

In this paper, we introduce and address a general formulation of Bayesian active learning with noisy observations that we call the *Equivalence Class Determination* problem. We show that, perhaps surprisingly, generalized binary search, and some other popular algorithms, perform poorly in this setting (i.e., in simulations where the true hypothesis is known, the algorithms do not always discover the true hypothesis quickly and reliably).

However, we show that a myopic (greedy) active learning algorithm using the **E**quivalence **C**lass **E**dge **C**utting (EC^2) objective function, can be proved to have an expected cost that is a multiple of the optimal test sequence cost. The key insight is that the EC^2 objective function for measuring information value satisfies *adaptive submodularity* (Golovin and Krause (2011)). Adaptive submodularity is a natural diminishing returns property that generalizes the classical notion of submodularity to adaptive policies, and leads to a provable performance guarantee as well as a shortcut in evaluating test choices (called "lazy evaluation").

Here are some possible *advantages* of BROAD approaches:

○ The posterior distribution of all theory and parameter probabilities is quickly recomputed for each subject after each trial (because it is a necessary step in finding the optimal next question). As a result, when the experiment is over, much of the data analysis is already done.

○ Because of that instant probability updating, BROAD creates a statistical parametric group portrait after each trial. The portraits show who seems most impatient, most averse to risk, most reciprocal, etc. These data could then be used to instantly form groups of different types of people to see how those differences affect group or market behavior.

○ Since BROAD techniques economize on information gained per minute, they are especially useful for subject pools who have scarce time or become bored or habituated quickly. Such groups include highly-trained professionals, internet subjects who can quit, human groups such as lesion patients or children, and animals that typically make long sequences of lab choices.[7]

○ The fact that the BROAD procedure generates sequences of test questions that are provably near-optimal can sharpen discourse about what different experimental designs are good and bad for. Novel BROAD designs which are unconventional should gain credibility if they have desirable informational properties. BROAD methods can also be used to judge the quality of older conventional designs[8].

○ At each trial it is easy to compute how much extra parameter precision (or model selection accuracy) is expected from asking one more question. This feature permits a cost-benefit method for optimally stopping

active learning (Nowak (2009); Balcan *et al.* (2006)), the results depend on the assumption that the same query can be made multiple times with independent noise. This is not appropriate in our setting; if a test subject makes a mistake in evaluating a pair of gambles, presenting the exact same choice later on is likely to elicit the same mistake, which induces correlation in the noise across tests.

[7] A sensible argument is that ethical treatment of subjects requires using as little of their time as possible to learn the most McClelland (1997); BROAD fits that bill.

[8] For example, retrospective analysis of influential psychology experiments on memory retention indicates that some early designs were remarkably near-optimal but others were not (Myung and Pitt (2009))

the experiment (e.g., Swartz and Choi (2009)), if information benefit and cost can be compared.

In economic choice applications, there is one possible imperfection in BROAD methods: In theory, subjects might prefer to strategically manipulate their early responses in order to get "better" (more economically valuable) future test questions. We pay special attention to the problem of manipulation, and discuss tests to detect it and methods to minimize it, within our experimental procedure. This is an important contribution of this paper, since strategic manipulation is not discussed in any detail in the most common applications (e.g. in CAT psychometric testing), and is a problem economists are well-equipped to address.

2.1 A brief history of related adaptive methods

The idea of optimizing an experimental design to measure parameters or test theories originated with Charles S. Peirce (Peirce (1967)), but did not have immediate influence. Dynamic design ideas began with Wald's influential sequential probability ratio test (Wald (1949)). Later contributions were made by Kiefer (1959), Atkinson and Fedorov (1975), and reviewed by Atkinson and Donev (1992) and Chaloner and Verdinelli (1995). D-optimal designs are optimal for estimating parameter values for a single theory. T-optimal designs optimally discriminate different theories and are more computationally challenging (cf. "landscaping", by D.J. Navarro and Myung (2004)). Many of the principles originally developed under D- and T-optimality now permeate standard experimental design practice.

The idea of adaptively customizing test choices goes back 100 years to Binet and Simon (1904). It was pushed forward decades later by the US Army (Bayroff *et al.* (1960)), and in psychometrics (Lord (1974); Weiss (1975); Weiss and Betz (1973)) in the 1960s and 70s. The most common method is computerized adaptive testing (CAT) for psychometric, intelligence (e.g., GRE, ASVAB military), health (PROMIS: Gershon *et al.* (2010)), and psychometric measurement (see Van Der Linden and Glas (2000); Wainer (1990)). CAT chooses maximally-informative items on a psychometric test, based on a subject's previous responses, using item response theory (IRT). There are now many applications in personnel selection and professional testing[9].

Besides their statistical advantage, there is also some evidence that adaptive methods improve test-taking motivation (Weiss and Betz (1973)), reduce boredom (Chapman *et al.* (2009), p. 3) and reduce differences across ethnic groups (Pine *et al.* (1979); Pine and Weiss (1978)).

Adaptive methods have also been used in neurophysiology (Lewi *et al.* (2009)), psychophysics (Kujala and Lukka (2006); L.A. Lesmes and Dosher (2006)), and medicine (Muller *et al.* (2007)). In marketing,

[9]E.g., ASVAB army; GRE and TOEFL; Microsoft Certification; see more examples at a clearinghouse at http://www.psych.umn.edu/psylabs/catcentral. Note that there are usually practical constraints imposed on large-scale CAT testing. These include balancing test content and test types and avoiding "overexposure" of highly diagnostic questions which can leak out across test-takers over time and undermine their future diagnostic value

adaptive choice-based "conjoint measurement" (ACBC) is used to discover which feature configurations of products consumers like most (e.g., Sawtooth software (Johnson *et al.* (2003)). Chapman *et al.* (2009) found modest success for ACBC in predicting actual product shares. A newer constraint satisfaction method called "poly-Q" takes a different approach, finding constrained feature weights which can explain previous choices (Abernethy *et al.* (2007); Toubia *et al.* (2004)). Note that improved methods will undoubtedly improve the ability to predict real behavior from responses which are hypothetical (often necessarily so). For example, Ding *et al.* (2005), Ding (2007), and Dong *et al.* (2010) describe incentive-aligned conjoint procedures. These methods use hypothetical data from conjoint measurements or product ranks combined with random selection of a single real product for purchase (based on WTP or rank inferred from conjoint or ranks), to boost incentive-compatibility of the conjoint responses. In cognitive psychology, adaptive methods have been championed by Myung and Pitt (2009), Myung *et al.* (2009), Cavagnaro *et al.* (2011), Cavagnaro *et al.* (2010, ming).

Two applications closely related to ours are Wang *et al.* (2010b); Toubia *et al.* (ming). Wang *et al.* (2010b) chooses questions that maximize Information-Gain, and Toubia *et al.* (ming) picks questions that maximize the norm of the Hessian of the posterior distribution at a point estimate (MLE) of the mode. None of the method have theoretical performance guarantees, and our paper provides the first theoretical bounds when responses are noisy. Furthermore, we use an accelerated greedy implementation of EC^2 that generates new questions in real-time, without the need to store the question sequence in advance, thus there are no practical limitations on the number of rounds of questions asked. Finally, we thoroughly consider the problem of strategic manipulation.

2.2 Models

In this section we describe six models of subjective valuation of risky choices; these models will later be tested experimentally using the EC^2 method. The models are presented first in order to ground discussion of the EC^2 method in a concrete application. Note that EC^2 can be applied to many different comparison of social science theories (e.g., time preference, social preference, etc.). The models are widely used in different areas of social sciences and in biology, and summarized in Table 2.1 for easy reference.

Choices are risky lotteries with probabilistic outcomes (a.k.a. prospects). The distribution of the payoffs in *lottery* L is a random variable with a range of *payoffs* $\mathcal{L} \equiv \{\ell_1, \ldots, \ell_k\}$. Lottery L is a distribution of objectively known probabilities p_i over payoffs, $p_i \equiv \mathbb{P}[L = \ell_i]$ (with $\sum_i^k p_i = 1$). For present purposes, a payoff is an amount of money gained or lost. However, in general the lotteries can have any payoffs which are subjectively valued by numerical utilities, or by other quantities (e.g., biological fitness).

Theory	Functional Form	Parameters
EV	$U_{EV}(L) = \mathbb{E}[L]$	none
CRRA	$\mathrm{U}_{CRRA}(L) = \begin{cases} \sum_i p_i (w+\ell_i)^{1-\theta}/(1-\theta) & \text{if } \theta \neq 1 \\ \sum_i p_i \log(w+\ell_i) & \text{if } \theta = 1 \end{cases}$	θ
PT	$\mathrm{U}_{PT}(L) = \sum_i v(\ell_i)\pi(p_i)$ where $v(\ell_i) = \begin{cases} \ell_i^{\rho} & \text{if } \ell_i \geq 0 \\ -\lambda(-\ell_i)^{\rho} & \text{if } \ell_i < 0 \end{cases}$	$\Theta_{PT} = \{\rho, \lambda, \alpha\}$
CPT	$\mathrm{U}_{CPT}(L) = \sum_{i=1}^{k} \pi_i^{+} v(\ell_i) + \sum_{j=k+1}^{n} \pi_j^{-} v(\ell_j)$(2.2.1)	$\Theta_{CPT} = \{\rho, \lambda, \alpha\}$
MVS	$\mathrm{U}_{MVS}(L) = w_\mu \mu - w_{\sigma^2}\sigma^2 + w_\nu \nu$(2.2.2)	$\Theta_{MVS} = \{w_\mu, w_{\sigma^2}, w_\nu\}$
NMVS	$\mathrm{U}_{NMVS}(L) = w_\mu \mu - w_\sigma \sigma + w_{\tilde{\nu}} \tilde{\nu}$(2.2.3)	$\Theta_{NMVS} = \{w_\mu, w_\sigma, w_{\tilde{\nu}}\}$

Table 2.1: We compare 6 models of risky choice, with no parameters: Expected Value (EV), one parameter: Constant Relative Risk Aversion (CRRA), adn 3 parameters: Prospect Theory (PT), Cumulative Prospect Theory (CPT), Mean-Variance-Skewness (MVS), Normalized Mean-Variance-Skewness

In expected utility (EU) theory, a lottery is valued by the probability-weighted sum of its possible subjective payoff utilities, $\sum_i^k p_i U(\ell_i)$ (von Neumann and Morgenstern (1947),Schoemaker (1982), Binmore (2009)).

Many different utility functions $U(\ell_i)$ have been applied and tested. We consider two popular functions $U(\ell_i)$: *Constant-Relative-Risk-Aversion (CRRA)*, whose functional form is given in Table 2.1. Wealth w is non-negative, and in subsequent experiments we set wealth to the subject's initial endowment[10].

The CRRA model is used most widely for modeling risky choices (Wakker (2010)), in economics (Holt and Laury (2005), Palacios-Huerta and Serrano (2006)), psychology (Luce and Krumhansl (1988)), and health (Bleichrodt *et al.* (1999)).

The restriction of CRRA to $\theta = 0$ gives *Expected Value (EV)*. Under EV decision-makers are neutral toward risk - they care only about the expectation of payoff. This is an unlikely general description (especially for large payoffs) but it is a useful benchmark.

[10]Wealth, w can be set somewhat arbitrarily by assuming a level of wealth for the population. In our framework, wealth can be treated as a parameter that can vary across subjects.

Prospect theory (PT) (Table 2.1) is a descriptive model of risky choice proposed as a more psychophysically-realistic alternative to EU (Kahneman and Tversky (1979)). The key difference is that values are thought to depend on gains and losses relative to a reference point (much as sensory evaluations are generally sensitive to short-term contrast effects with previous stimuli and experiences). Reference-dependence is important because of two other properties assumed in prospect theory: First, diminishing marginal sensitivity is assumed to hold both as gains increase in magnitude (from a zero-gain point of reference) *and* as losses increase in magnitude (i.e., become more negative). This assumption implies concavity of gain utility and convexity of loss disutility (contrary to the typical EU analysis in economics, which assumes payoffs are integrated wealth positions). Second, losses are assumed to loom larger than gains, so that $v(\ell_i) < -v(-\ell_i)$ for $\ell_i > 0$.

Besides reference-dependence, PT differs from simple EU by assuming that objective probabilities are transformed into a nonlinear subjective decision weighting functions. These decision weights $\pi(p_i)$ are then used to weight reference-adjusted payoffs.[11]. As in many other experiments, we assume a reference outcome at 0. That is, positive outcomes are *gains* and negative outcomes are *losses* in all choices, relative to a benchmark of a subject's wealth at the beginning of the experiment.

A CRRA power function is used for the value function v and losses are weighed by a factor λ (typically greater than one); The functional form of the prospect value of a lottery is shown in Table 2.1. Several decision weight functions $\pi(p_i)$ have been proposed (see Hsu *et al.* (2009), Booij *et al.* (2010)). In order to limit the parametric features being compared across models, we consider only the one-parameter decision weight function axiomatically derived in Prelec (1998), which is $\pi(p_i) = (e^{-(\log(1/p_i))^\alpha})$. The parameter α expresses the degree of inflection: $\alpha = 1$ is linear weighting (as in EU); $\alpha < 1$ expresses the inflection hypothesized in prospect theory (Kahneman and Tversky (1979)), overweighting probabilities below $p = 1/e$ and underweighting higher probabilities. The parameters $\Theta_{PT} = \{\rho, \lambda, \alpha\}$ therefore represent utility nonlinearity, loss aversion and decision weight inflection, respectively.

In the PT model just described, decision weights of all probabilities do not generally add up to one. It is well-known that this property can generate predicted violations of first-order stochastic dominance. In Tversky and Kahneman (1992) the original PT is extended by using rank-dependent weights (following Quiggin (1982)), in a form called *Cumulative Prospect Theory (CPT)*. In CPT, a payoff is weighted by the difference between its transformed decumulative probability and the decumulative probability of the payoff ranked just below it in magnitude. Consider a gamble with two positive gain payoffs $\ell_1 > \ell_2$, with objective probabilities p_1 and p_2 and a zero payoff. Then the ranked payoffs are weighted $\pi(p_1), \pi(p_2 + p_1) - \pi(p_1)$

[11] Evidence for empirical accuracy of prospect theory has come from many methods and domains: Lab experiments (e.g., Harless and Camerer (1994), Hsu *et al.* (2009)); field surveys with large representative samples (Booij *et al.* (2010)); pricing of stocks (Benartzi and Thaler (1995)), and options (Gurevich *et al.* (2009), Barberis *et al.* (2001), Heath *et al.* (1999)); game shows (Post et al); and neuroscience (Trepel *et al.* (2005), Fox and Poldrack (2009), Christopoulos *et al.* (2009), Hsu *et al.* (2009), De Martino *et al.* (2006b), Tom *et al.* (2007))

and $1 - \pi(p_2 + p_1)$.

More generally, rank payoffs from 1 (best) to n (worst) by $\ell_1 \geq \cdots \geq \ell_k \geq 0 \geq \ell_k + 1 \geq \cdots \geq \ell_n$. The first k payoffs are gains and payoffs ranked $k+1$ to n are losses. Denote decision weight in the gain and loss domains by $w^+(p)$ and w^-. Then the weights are

$$\begin{aligned} \text{for } i \leq k : \pi_i^+ &= w^+(p_i + \cdots + p_1) - w^+(p_{i-1} + \cdots + p_1) \\ \text{for } j > k : \pi_j^- &= w^-(p_j + \cdots + p_n) - w^-(p_{j+1} + \cdots + p_n) \end{aligned}$$

Then the prospect value depends on ranks and the rank dependence of decision weights (refer to Table 2.1).

Besides EU variants, PT and CPT, another class of models assumes that risky choices are valued by computing and combining their statistical moments– their mean, variance, and skewness (and possibly higher moments too). Moments-based models have been most widely used in financial economics (Sharpe (1964); Markowitz (1968)) and biology[12].

Two models defined next capture the most common types of moments-based models. *Mean-Variance-Skewness (MVS)*, in Table 2.1, has parameters $\Theta_{MVS} = \{w_\mu, w_{\sigma^2}, w_\nu\}$ are the weights for the mean, variance and skewness of the lottery respectively, with $w_\mu, w_{\sigma^2}, w_\nu \geq 0$.

Since variance and skewness become much larger when payoffs are scaled up, it is also reasonable to include a standardized version of weighted moments that scales with payoff increases more gracefully, which gives the *Normalized Mean-Variance-Skewness (NMVS)* (Table 2.1), where parameters $\Theta_{SMVS} = \{w_\mu, w_\sigma, w_{\tilde{\nu}}\}$ are the weights for the mean, standard deviation and standardized skewness of the lottery respectively, where $w_\mu, w_\sigma, w_{\tilde{\nu}} \geq 0$.

We therefore have six different models to compare. Many previous experiments tried to design choices which are most different (i.e., most diagnostic) in these different models. For example, EU and MVS approaches predict either a general aversion to payoff variance, or a preference. PT and CPT allow an aversion to variance over gains and a preference for variance over losses. It is therefore easy to design choices with gains and losses in which the theories make different predictions (see Wang *et al.* (2010b), Toubia *et al.* (ming)).

A subtler difference arises between PT/CPT, and MVS/SMVS approaches. PT and CPT predict a high subjective value for risks with low-probability high-payoff gains (due to nonlinear decision weights $\pi(p)$ on the extreme gains). MVS and SMVS also predict high subjective value for such risks because their skewness is positive. Distinguishing the two approaches therefore requires a deft touch, finding mixtures of high-

[12]Criticisms of the mean-variance approach emerged early in its evolution. EU valuations can be approximated by weighting mean and variance, using a Taylor expansion argument. However, mean-variance is strictly compatible with EU-based choice only if payoff utility is quadratic or if risky choice payoffs are normally distributed (Tobin (1958), Hanoch and Levy (1969)). Quadratic utility can lead to stochastic dominance valuations, and implies that absolute risk-aversion is increasing, which is unlikely (Pratt (1964) and Arrow (1971)). However, analyses suggest the conclusions from mean-variance models are robust to deviations from quadratic utility and normality (Levy and Markowitz (1979); Kroll *et al.* (1984)).

payoff amount and high-payoff (weighted) probability such that pure skewness preference and high-payoff overweighting can be distinguished.

2.3 Dynamic Optimal Designs

We begin with intuition about EC^2, then describe the method formally. It has three powerful properties which other information criteria do not have:

◦ Most methods are sensitive to noise.[13] The empirical applications which we have an mind all feature responses with some degree of stochastic noise (i.e., people choosing according to decision rule will not make the same choice every single time from a fixed choice set). Many of the good properties of previous adaptive methods rely on deterministic, no-noise processes (so that hypotheses are immediately eliminated by a single unpredicted "mistake"). Our method EC^2 reduces the noisy case to a formally tractable noiseless case, so good performance properties carry over while noise is allowed.

◦ It can be proved that there is a worst-case cost bound for greedy EC^2 which depends on $ln(1/p_{min})$ (where p_{min} is the prior probability of the least-likely theory). What does that imply? In a medical setting, for example, there could be a dangerous disease which is extremely rare (p_{min} is low) but highly contagious; and figuring out for sure whether a patient has the disease is necessary for protecting public health. This kind of diagnosis will take many tests in the worst case. However, in the empirical social science settings we are considering, the lowest p_{min} is rarely very low because it is usually adequate to consider only a modest number of possible theories and parameters (p_{min} might be on the order of 10^{-3}, for example, so that the worst-case cost multiplier is 6.9).

◦ Since our EC^2 information criteria is adaptively submodular, the marginal information value of a test can only go down after other tests have been done. This permits "lazy evaluation" which reduces runtime (Leskovec *et al.* (2007); Minoux (1978)).

To illustrate, suppose there are 30,000 possible tests. The procedure begins by computing the value, according to the EC^2 objective criteria (revealed in later sections), of all 30,000 tests. The highest-value test is conducted on trial 1, the result is recorded, and all model probabilities are updated. On trial 2, if a candidate new test has a computed value V, then all the tests with a value lower than V in the initial pre-trial 1 evaluation can be safely ignored (since those values may change after the trial 1 results, but can never become larger than V). This insight can be exploited, by considering tests in order of maximum previously computed values (Refer to Appendix for details).

Next we will provide some background on general methods with no noise, which sets the stage for EC^2. The notation is borrowed from computer science (see Golovin and Krause (2011)) and is designed to apply to a very wide range of domains (e.g., sensor management, medical testing, etc.). A simple visual illustration is given below (Figure 2) for the case of hypotheses about risky choice.

[13] Appendix B illustrates an example where information gain and value of information (VoI) perform badly.

2.3.1 Bayesian Active Learning in the Noiseless Case

In the Bayesian active learning problem, we would like to distinguish among a given set of hypotheses $\mathcal{H} = \{h_1, \ldots, h_n\}$ by performing tests from a set $\mathcal{T} = \{1, \ldots, N\}$ of possible tests. Running test t incurs a cost of $c(t)$ and produces an outcome from a finite set of outcomes $\mathcal{X} = \{1, 2, \ldots, \ell\}$.

In our setting, $\mathcal{H}$ would be the parameter values of models, and $\mathcal{T}$, the set of all possible choice pairs. $\mathcal{X}$ is the choice observed ($\mathcal{X} = \{0, 1\}$ for a choice between two lotteries).

We let H denote the random variable which equals the true hypothesis, and model the outcome of each test t by a random variable X_t taking values in $\mathcal{X}$. We denote the observed outcome of test t by x_t. We further suppose we have a prior distribution P modeling our assumptions on the joint probability $P(H, X_1, \ldots, X_N)$ over the hypotheses and test outcomes.

In the noiseless case, we assume that the outcome of each test is deterministic given the true hypothesis, i.e., for each $h \in \mathcal{H}$, $P(X_1, \ldots, X_N \mid H = h)$ is 0 or 1. Thus, each hypothesis h is associated with a particular vector of positive test outcomes. We assume that no two hypotheses lead to the same outcomes for all tests.[14] Therefore, if we perform all tests, we can uniquely determine the true hypothesis.

However in most applications we will wish to avoid performing every possible test, as this is prohibitively expensive. Our goal is to find an adaptive policy for running tests that allows us to determine the value of H with certainty while minimizing the cost of the tests performed. Formally, a policy π (also called a conditional plan) is a partial mapping π from partial observation vectors $\mathbf{x}_\mathcal{A}$ to tests, specifying which test to run next (or whether we should stop testing) for any observation vector $\mathbf{x}_\mathcal{A}$. Hereby, $\mathbf{x}_\mathcal{A} \in \mathcal{X}^\mathcal{A}$ is a vector of outcomes indexed by a set of tests $\mathcal{A} \subseteq \mathcal{T}$ that we have performed so far [15] (e.g., the set of binary choices between risky gamble pairs, or outcomes of a set of medical tests). After having made observations $\mathbf{x}_\mathcal{A}$, we can rule out inconsistent hypotheses. We denote the set of hypotheses consistent with an observation vector, or event Λ, by $\mathcal{V}(\Lambda) \equiv \{h \in \mathcal{H} : P(h \mid \Lambda) > 0\}$. ($\mathcal{V}(\Lambda)$ is often called the *version space* associated with Λ.)

We call a policy *feasible* if it is guaranteed to uniquely determine the correct hypothesis. That is, upon termination with observation $\mathbf{x}_\mathcal{A}$, it must hold that $|\mathcal{V}(\mathbf{x}_\mathcal{A})| = 1$. We can define the expected cost of a policy π by

$$c(\pi) \equiv \sum_h P(h) c(\mathcal{T}(\pi, h))$$

where $\mathcal{T}(\pi, h) \subseteq \mathcal{T}$ is the set of tests run by policy π in case $H = h$.

Our goal is to find a feasible policy π^* of minimum expected cost, i.e.,

$$\pi^* = \arg\min \{c(\pi) : \pi \text{ is feasible}\} \tag{2.3.1}$$

[14] That is, we implicitly assume that at least one test can be done that distinguishes any two hypotheses.

[15] Formally we also require that $(x_t)_{t \in \mathcal{B}} \in \mathrm{dom}(\pi)$ and $\mathcal{A} \subseteq \mathcal{B}$, implies $(x_t)_{t \in \mathcal{A}} \in \mathrm{dom}(\pi)$ (c.f., Golovin and Krause (2010b)).

A policy π can be naturally represented as a decision tree T^π, and thus problem (3.2.1) is often called the *Optimal Decision Tree* (ODT) problem.

Unfortunately, obtaining an approximate policy π for which $c(\pi) \leq c(\pi^*) \cdot o(\log(n))$ is NP-hard (Chakaravarthy *et al.* (2007)). Hence, various heuristics are employed to solve the Optimal Decision Tree problem and its variants. Two of the most popular heuristics are to select tests by greedily maximizing the *information gain* (IG) conditioned on previous test outcomes, and *generalized binary search* (GBS). (A search is called "greedy" if it only considers the immediate effect of tests on some fixed measure of progress).

Both of the IG and GBS heuristics are greedy and after having made observations $\mathbf{x}_\mathcal{A}$ will select

$$t^* = \arg\max_{t \in \mathcal{T}} \Delta_{\text{Alg}}(t \,|\, \mathbf{x}_\mathcal{A}) / c(t),$$

where Alg $\in \{\text{IG}, \text{GBS}\}$.

Here, $\Delta_{\text{IG}}(t \,|\, \mathbf{x}_\mathcal{A}) \equiv \mathbb{H}(\mathbf{X}_\mathcal{T} \mid \mathbf{x}_\mathcal{A}) - \mathbb{E}_{x_t \sim X_t | \mathbf{x}_\mathcal{A}}[\mathbb{H}(\mathbf{X}_\mathcal{T} | \mathbf{x}_\mathcal{A}, x_t)]$ is the marginal information gain measured with respect to the Shannon entropy $\mathbb{H}(\mathbf{X}) \equiv \mathbb{E}_\mathbf{x}[-\log_2 P(\mathbf{x})]$, and $\Delta_{\text{GBS}}(t \,|\, \mathbf{x}_\mathcal{A}) \equiv P(\mathcal{V}(\mathbf{x}_\mathcal{A})) - \sum_{x \in \mathcal{X}} P(X_t = x \mid \mathbf{x}_\mathcal{A}) P(\mathcal{V}(\mathbf{x}_\mathcal{A}, X_t = x))$ is the expected reduction in version space probability mass. Thus, both heuristics greedily choose the test that maximizes the benefit-cost ratio, measured with respect to their particular (informational) benefit functions. The heuristics stop after running a set of tests $\mathcal{A}$ such that $|\mathcal{V}(\mathbf{x}_\mathcal{A})| = 1$, i.e., once the true hypothesis has been uniquely determined. It turns out that for the (noiseless) Optimal Decision Tree problem, the IG and GBS heuristics are equivalent (Zheng *et al.* (2005)).

Interestingly, despite its greedy property GBS has been shown (Kosaraju *et al.* (1999), Dasgupta (2004), Guillory and Bilmes (2009), Golovin and Krause (2010b)) to obtain near-optimal expected cost: the strongest known bound is $c(\pi_{GBS}) \leq c(\pi^*)(\ln(1/p_{\min}) + 1)$ where $p_{\min} \equiv \min_{h \in \mathcal{H}} P(h)$.

Let $\mathbf{x}_S(h)$ be the unique vector $\mathbf{x}_S \in \mathcal{X}^S$ such that $P(\mathbf{x}_S \mid h) = 1$. The result above is proved by exploiting the fact that $f_{GBS}(S, h) \equiv 1 - P(\mathcal{V}(\mathbf{x}_S(h))) + P(h)$ is *adaptive submodular*[16] and *strongly adaptively monotone*, Golovin and Krause (2010b). These two properties will turn out to be crucial for proving theoretical guarantees of approximation bounds for EC^2 later.

Call $\mathbf{x}_\mathcal{A}$ a *subvector* of $\mathbf{x}_\mathcal{B}$ if $\mathcal{A} \subseteq \mathcal{B}$ and $P(\mathbf{x}_\mathcal{B} \mid \mathbf{x}_\mathcal{A}) > 0$. In this case we write $\mathbf{x}_\mathcal{A} \prec \mathbf{x}_\mathcal{B}$. A function $f : 2^\mathcal{T} \times \mathcal{H}$ is called adaptive submodular w.r.t. a distribution P, if for any $\mathbf{x}_\mathcal{A} \prec \mathbf{x}_\mathcal{B}$ and any test t it holds that

$\Delta(t \,|\, \mathbf{x}_\mathcal{A}) \geq \Delta(t \,|\, \mathbf{x}_\mathcal{B})$, where

$$\Delta(t \,|\, \mathbf{x}_\mathcal{A}) \equiv \mathbb{E}_H[f(\mathcal{A} \cup \{t\}, H) - f(\mathcal{A}, H) \;\mid \mathbf{x}_\mathcal{A}].$$

[16] Submodularity has proved useful in many other economic contexts, e.g. Chambers and Echenique (2008) and Shirai (2010)

Thus, f is adaptive submodular if the expected marginal benefits $\Delta(t \,|\, \mathbf{x}_{\mathcal{A}})$ of adding a new test t can only decrease as we gather more observations. f is called *strongly adaptively monotone* w.r.t. P if, informally, "more observations never hurt" with respect to the expected reward. Formally, for all $\mathcal{A}$, all $t \notin \mathcal{A}$, and all $x \in \mathcal{X}$ we require

$$\mathbb{E}_H[f(\mathcal{A}, H) \mid \mathbf{x}_{\mathcal{A}}] \leq \mathbb{E}_H[f(\mathcal{A} \cup \{t\}, H) \mid \mathbf{x}_{\mathcal{A}}, X_t = x].$$

The performance guarantee for IG and GBS follows from the following general result about the greedy algorithm for adaptive submodular functions (applied with $Q = 1$ and $\eta = p_{\min}$):

Theorem 1 (Theorem 10 of Golovin and Krause (2010b) with $\alpha = 1$)**.** *Suppose $f : 2^{\mathcal{T}} \times \mathcal{H} \to \mathbb{R}_{\geq 0}$ is adaptive submodular and strongly adaptively monotone with respect to P and there exists Q such that $f(\mathcal{T}, h) = Q$ for all h. Let η be any value such that $f(S, h) > Q - \eta$ implies $f(S, h) = Q$ for all sets S and hypotheses h.*

Then for self–certifying[17] *instances the adaptive greedy policy π satisfies*

$$c(\pi) \leq c(\pi^*) \left(\ln\left(\frac{Q}{\eta}\right) + 1\right).$$

In the following sections, we will use the concept of adaptive submodularity to provide the first approximation guarantees of test cost for Bayesian active learning with noisy observations. The concrete form of the objective criteria f is discussed in section 2.4, along with an effective implementation.

2.3.2 The Equivalence Class Determination Problem

There are many ways to include and estimate noise in choice. The most popular method in social science is to estimate some functional specification for the noise term, and estimate the amount of noise as a parameter(s) in the general model of choice[18]. For adaptive design this approach has been pursued by Wang, Filiba and Camerer (2011). This can be problematic for two reasons:

First, with finite test vectors, it will generally be impossible to guarantee feasibility– i.e., the property that a test method will certainly locate the true hypothesis eventually– which undermines approximation guarantees. Put differently, the *version space* will not shrink rapidly because noise means that it is difficult to conclusively eliminate hypotheses.

Second, honing in on likely models both accurately and precisely requires simultaneously identifying the noise parameter and other parameters whose identification depends on noise. The results in Wang et al. (2011)

[17] The technical requirement that instances be *self-certifying* means that the policy will have proof that it has obtained the maximum possible objective value, Q, immediately upon doing so. It is not difficult to show that this is the case with the instances we consider in this paper. See Golovin and Krause (2010b) for more detail.

[18] See Becker *et al.* (1963), Loomes and Sugden (1995), Loomes and Sugden (1998), Loomes *et al.* (2002), Moffatt (2005), Moffatt and Peters (2001), Hey (2005), Hey and Orme (1994), Wilcox (2008)

are somewhat pessimistic in this regard (i.e., precision in identifying the amount of noise does not increase much over dozens of experimental trials.)

We therefore pursue a different approach. The trick is to model noise by assuming that there is random switching of "true" deterministic responses. To gain intuition, consider a simple model where tests have binary outcomes, and we know that the outcome of exactly one test, chosen uniformly at random unbeknownst to us, is flipped. If any pair of hypotheses $h \neq h'$ differs by the outcome of at least three tests, we can still uniquely determine the correct hypothesis after running all tests. In this case we can reduce the noisy active learning problem to the noiseless setting by creating N "noise-injected" copies for each hypothesis. Each noisy copy is obtained by flipping the outcome of one of the N tests. The modified prior P' would then assign mass $P'(h') = P(h)/N$ to each noisy copy h' of h. The conditional distribution $P'(\mathbf{X}_{\mathcal{T}} \mid h')$ is still deterministic (since flipping the outcome of one of the tests changes it from 0 to 1 or vice versa). Thus, each deterministic hypothesis h_i in the original problem is now associated with a set $\mathcal{H}_i$ of deterministic hypotheses in the modified problem. However, instead of selecting tests to determine exactly which noisy copy has been realized, we only care which set of $\mathcal{H}_i$ (all generated by flipping elements of h_i) is realized. That is, we will not necessarily be able to locate the correct noisy copy, but we can locate the family $\mathcal{H}_i$ of noisy copies, which all spring from a single hypothesis h_i. Note that until we have enough observations, multiple hypotheses h_i would exist in the same equivalence class $\mathcal{H}_i$.

In the *Equivalence Class Determination problem*[19], a set of hypotheses $\mathcal{H}$ is partitioned into a set of m equivalence classes $\{\mathcal{H}_1, \ldots, \mathcal{H}_m\}$ so that $\mathcal{H} = \biguplus_{i=1}^{m} \mathcal{H}_i$. The goal is to determine which class $\mathcal{H}_i$ the true hypothesis lies in.

Formally, upon termination with observations $\mathbf{x}_{\mathcal{A}}$ we require that $\mathcal{V}(\mathbf{x}_{\mathcal{A}}) \subseteq \mathcal{H}_i$ for some i. As with the ODT problem, the goal is to minimize the expected cost of the tests, where the expectation is taken over the true hypothesis sampled from P. In §2.3.3, we will show how the ECD problem arises naturally from Bayesian experimental design problems in probabilistic models.

Given the fact that GBS performs near-optimally on the Optimal Decision Tree problem, a natural approach to solving ECD would be to run GBS until the termination condition is met. Unfortunately, and perhaps surprisingly, GBS can perform very poorly on the ECD problem. Consider an instance with a uniform prior over n hypotheses, $h_1, \ldots, h_n$, and two equivalence classes $\mathcal{H}_1 \equiv \{h_i : 1 \leq i < n\}$ and $\mathcal{H}_2 \equiv \{h_n\}$. There are tests $\mathcal{T} = \{1, \ldots, n\}$ such that $h_i(t) = \mathbf{1}[i = t]$, all of unit cost. Hereby, $\mathbf{1}[\Lambda]$ is the indicator variable for event Λ. In this case, the optimal policy only needs to select test n. However GBS may select tests $1, 2, \ldots, n$ in order until running test t, where $H = h_t$ is the true hypothesis. Given our uniform prior, it takes $n/2$ tests

[19] Bellala *et al.* (2010) simultaneously studied ECD, and, like us, used it to model active learning with noise (Bellala *et al.* (2009)). They developed an extension of GBS for ECD. We defer a detailed comparison of our approaches to future work.

in expectation until this happens, so that GBS costs, in expectation, $n/2$ times the optimal expected cost in this instance.

The poor performance of GBS in this instance may be attributed to its lack of consideration for the equivalence classes. Another natural heuristic would be to run the greedy IG policy, only with the entropy measured with respect to the probability distribution on *equivalence classes* rather than hypotheses. Call this policy π_{IG}. It is clearly aware of the equivalence classes, as it adaptively and myopically selects tests to reduce the uncertainty of the realized class, measured w.r.t. the Shannon entropy. However, we can show there are instances in which it costs $\Omega(n/\log(n))$ times the optimal cost, even under a uniform prior (see Appendix B for an example).

The EC^2 selection criterion. The reason why GBS fails is because reducing the version space mass does not necessarily facilitate differentiation among the classes $\mathcal{H}_i$. The reason why IG fails is that there are complementarities among tests; a set of tests can be far better than the sum of its parts (so greedy test choice can fail badly because it only considers one test at a time).

Thus, we would like to optimize an objective function that encourages differentiation among classes, but lacks complementarities. We adopt an elegant idea from Dasgupta (2006), and define weighted edges between hypotheses that we would like distinguish between. However, instead of introducing edges between arbitrary pairs of hypotheses (as done in Dasgupta (2006)), we only introduce edges between hypotheses in different classes[20]. Tests will allow us to cut edges inconsistent with their outcomes, and we aim to eliminate all inconsistent edges while minimizing the expected cost incurred. We now formalize this intuition.

Specifically, we define a set of edges $\mathcal{E} = \cup_{1\leq i<j\leq m} \{\{h, h'\} : h \in \mathcal{H}_i, h' \in \mathcal{H}_j\}$, consisting of all (unordered) pairs of hypotheses belonging to distinct classes. These are the edges that must be *cut*, by which we mean for any edge $\{h, h'\} \in \mathcal{E}$, at least one hypothesis in $\{h, h'\}$ must be ruled out (i.e., eliminated from the version space). Hence, a test t run under true hypothesis h is said to cut edges $\mathcal{E}_t(h) \equiv \{\{h', h''\} : h'(t) \neq h(t) \text{ or } h''(t) \neq h(t)\}$. See Fig. 3.1 for an illustration. We define a weight function $w : \mathcal{E} \to \mathbb{R}_{\geq 0}$ by $w(\{h, h'\}) \equiv P(h) \cdot P(h')$. We extend the weight function to an additive (modular) function on sets of edges in the natural manner, i.e., $w(\mathcal{E}') \equiv \sum_{e\in\mathcal{E}'} w(e)$. The objective f_{EC} that we will greedily maximize is then defined as the weight of the edges cut (EC):

$$f_{EC}(\mathcal{A}, h) \equiv w\Big(\bigcup_{t\in\mathcal{A}} \mathcal{E}_t(h)\Big) \tag{2.3.2}$$

The key insight that allows us to prove approximation guarantees for f_{EC} is that f_{EC} shares the same beneficial properties that make f_{GBS} amenable to efficient greedy optimization.

[20]We do not care to distinguish noisy copies within an equivalence class, so there is no need to create edges between these copies.

Proposition 2. *The objective f_{EC} is strongly adaptively monotone and adaptively submodular.*

Based on the objective f_{EC}, we can calculate the marginal benefits for test t upon observations $\mathbf{x}_{\mathcal{A}}$ as

$$\Delta_{\text{EC}}(t \,|\, \mathbf{x}_{\mathcal{A}}) \equiv \mathbb{E}_H[f_{EC}(\mathcal{A} \cup \{t\}, H) - f_{EC}(\mathcal{A}, H) \;\; | \; \mathbf{x}_{\mathcal{A}}] \,.$$

We call the adaptive policy π_{EC} that, after observing $\mathbf{x}_{\mathcal{A}}$, greedily selects test $t^* \in \arg\max_t \Delta_{\text{EC}}(t \,|\, \mathbf{x}_{\mathcal{A}}) / c(t)$, the EC^2 algorithm (for *equivalence class edge cutting*).

Note that these instances are self–certifying, because we obtain maximum objective value if and only if the version space lies within an equivalence class, and the policy can certify this condition when it holds. So we can apply Theorem 1 to show EC^2 obtains a $\ln(Q/\eta) + 1$ approximation to Equivalence Class Determination. Hereby, $Q = w(\mathcal{E}) = 1 - \sum_i (P(h \in \mathcal{H}_i))^2 \leq 1$ is the total weight of all edges that need to be cut, and $\eta = \min_{e \in \mathcal{E}} w(e) \geq p_{\min}^2$ is a bound on the minimum weight among all edges. We have the following result:

Theorem 3. *Suppose $P(h)$ is rational for all $h \in \mathcal{H}$. For the adaptive greedy policy π_{EC} implemented by* EC^2 *it holds that*

$$c(\pi_{EC}) \leq (2 \ln(1/p_{\min}) + 1) c(\pi^*),$$

where $p_{\min} \equiv \min_{h \in \mathcal{H}} P(h)$ is the minimum prior probability of any hypothesis, and π^* is the optimal policy for the Equivalence Class Determination problem.

2.3.3 Noisy Observations

In this section we address the case of noisy observations, using ideas from §2.3.2, and generalize the notion of the noise model in previous sections. With noisy observations, the conditional distribution $P(X_1, \ldots, X_N \mid h)$ is no longer deterministic. We model the noise using an additional random variable Θ. The vector of test outcomes $\mathbf{x}_{\mathcal{T}}$ is assumed to be an arbitrary, deterministic function $\mathbf{x}_{\mathcal{T}} : \mathcal{H} \times \text{supp}(\Theta) \to \mathcal{X}^N$; hence $\mathbf{X}_{\mathcal{T}} \mid h$ is distributed as $\mathbf{x}_{\mathcal{T}}(h, \Theta_h)$ where Θ_h is distributed as $P(\theta \mid h)$. For example, there might be up to $s = |\text{supp}(\Theta)|$ ways any particular disease could manifest itself, with different patients with the same disease suffering from different symptoms.

In cases where it is always possible to identify the true hypothesis, i.e., $\mathbf{x}_{\mathcal{T}}(h, \theta) \neq \mathbf{x}_{\mathcal{T}}(h', \theta')$ for all $h \neq h'$ and all $\theta, \theta' \in \text{supp}(\Theta)$, we can reduce the problem to Equivalence Class Determination with hypotheses $\{\mathbf{x}_{\mathcal{T}}(h, \theta) : h \in \mathcal{H}, \theta \in \text{supp}(\Theta)\}$ and equivalence classes $\mathcal{H}_i \equiv \{\mathbf{x}_{\mathcal{T}}(h_i, \theta) : \theta \in \text{supp}(\Theta)\}$ for all i. Then Theorem 3 immediately yields that the approximation factor of EC^2 is at most $2 \ln(1/\min_{h,\theta} P(h, \theta)) + 1$, where the minimum is taken over all (h, θ) in the support of P. In the unit cost case, running EC^2 with a modified prior à la Kosaraju *et al.* (1999), allows us to obtain an $\mathcal{O}(\log |\mathcal{H}| + \log |\text{supp}(\Theta)|)$ approximation

factor. Note this model allows us to incorporate complex correlations in the noise terms.

However, a major challenge when dealing with noisy observations is that it is not always possible to distinguish distinct hypotheses. Even after we have run all tests, there will generally still be uncertainty about the true hypothesis, i.e., the posterior distribution $P(H \mid \mathbf{x}_{\mathcal{T}})$ obtained using Bayes' rule may still assign non-zero probability to more than one hypothesis. If so, uniquely determining the true hypothesis is not possible. Instead, we imagine that there is a set $\mathcal{D}$ of possible *decisions* we may make after (adaptively) selecting a set of tests to perform and we must choose one (e.g., we must decide how to treat the medical patient, which scientific theory to adopt, or which classifier to use, given our observations). Thus our goal is to gather data to make effective decisions (Howard (1966)). Formally, for any decision $d \in \mathcal{D}$ we take, and each realized hypothesis h, we incur some loss $\ell(d, h)$. Decision theory recommends, after observing $\mathbf{x}_{\mathcal{A}}$, to choose the decision d^* that minimizes the *risk*, i.e., the expected loss, namely $d^* \in \arg\min_d \mathbb{E}_H[\ell(d, H) \mid \mathbf{x}_{\mathcal{A}}]$.

A natural goal in Bayesian active learning is thus to adaptively pick observations, until we are guaranteed to make the same decision (and thus incur the same expected loss) that we would have made had we run *all* tests. Thus, we can reduce the noisy Bayesian active learning problem to the ECD problem by defining the equivalence classes over all test outcomes that lead to the same minimum risk decision. Hence, for each decision $d \in \mathcal{D}$, we define

$$\mathcal{H}_d \equiv \{\mathbf{x}_{\mathcal{T}} : d = \arg\min_{d'} \mathbb{E}_H[\ell(d', H) \mid \mathbf{x}_{\mathcal{T}}]\}. \tag{2.3.3}$$

If multiple decisions minimize the risk for a particular $\mathbf{x}_{\mathcal{T}}$, we break ties arbitrarily. Identifying the best decision $d \in \mathcal{D}$ then amounts to identifying which equivalence class $\mathcal{H}_d$ contains the realized vector of outcomes, which is an instance of ECD.

One common approach to this problem is to myopically pick tests maximizing the decision-theoretic *value of information* (VoI) (Howard (1966)): $\Delta_{\text{VoI}}(t \,|\, \mathbf{x}_{\mathcal{A}}) \equiv \min_d \mathbb{E}_H[\ell(d, H) \mid \mathbf{x}_{\mathcal{A}}] - \mathbb{E}_{x_t \sim X_t | \mathbf{x}_{\mathcal{A}}}[\min_d \mathbb{E}_H[\ell(d, H) \mid \mathbf{x}_{\mathcal{A}}, x_t]]$. The VoI of a test t is the expected reduction in the expected loss of the best decision due to the observation of x_t. However, we can show there are instances in which such a policy pays $\Omega(n/\log(n))$ times the optimal cost, even under a uniform prior on (h, θ) and with $|\operatorname{supp}(\Theta)| = 2$ (Appendix B contains an example). In contrast, on such instances our algorithm obtains an $\mathcal{O}(\log n)$ approximation. More generally, we have the following result for EC^2 as an immediate consequence of Theorem 3.

Theorem 4. *Fix hypotheses $\mathcal{H}$, tests $\mathcal{T}$ with costs $c(t)$ and outcomes in $\mathcal{X}$, decision set $\mathcal{D}$, and loss function ℓ. Fix a prior $P(H, \Theta)$ and a function $\mathbf{x}_{\mathcal{T}} : \mathcal{H} \times \operatorname{supp}(\Theta) \to \mathcal{X}^N$ which define the probabilistic noise model. Let $c(\pi)$ denote the expected cost of π incurs to find the best decision, i.e., to identify which equivalence class $\mathcal{H}_d$ the outcome vector $\mathbf{x}_{\mathcal{T}}$ belongs to. Let π^* denote the policy minimizing $c(\cdot)$, and let π_{EC} denote the adaptive*

policy implemented by EC^2. *Then it holds that*

$$c(\pi_{EC}) \leq \left(2\ln\left(\frac{1}{p'_{\min}}\right)+1\right)c(\pi^*),$$

where $p'_{\min} \equiv \min_{h\in\mathcal{H}}\{P(h,\theta)\ :\ P(h,\theta)>0\}$.

If all tests have unit cost, by using a modified prior (Kosaraju *et al.* (1999)) the approximation factor can be improved to $\mathcal{O}\left(\log|\mathcal{H}|+\log|\operatorname{supp}(\Theta)|\right)$ as in the case of Theorem 3.

2.3.4 A simple illustration

Intuitively, EC^2 chooses the test which (a) distinguishes two hypotheses, connected by an edge, which make different predictions on that test, and (b) has the largest product of those two hypotheses' probability. The method is like a sports tournament in which all pairs of challengers that make different predictions are created (edges). Pairs are ranked by the product of their probabilities of being the "best" (i.e., most likely). The pair of contending theories with the highest probability product are then chosen to compete on a conclusive test (which will knock out one and be consistent with the other; In fact, a challenge question can maximize the sum of probability products for multiple edges, hence knocking out multiple theories at the same time). The idea is to choose the challenge which knocks out one of two theories –in our context, one of two theories that have the highest product of likelihoods of being true– as quickly as possible.

Figures 2.1 and 3.1 show a graphic illustration of how EC^2 works. 2.1 (top left) shows an equivalence-class: Each oval represents an equivalence class (EC) of theory predictions. The theory predicts choices of [0,0,0] for three tests X_1, X_2, X_3. A "noisy copy" of the theory's prediction [1*,0,0] is randomly generated by flipping the first choice from 0 to 1. (No noisy copy is created for the bottom oval theory.) If the noise rate is 10%, then the probability of observing choices [1*,0,0] is 0.1*0.9*0.9 times the probability of the true predicted choices [0,0,0]. 2.1 (top right) illustrates equivalence-class edges: Edges connect all pairs of predictions between equivalence sets (although, importantly, *no edges* connect *within* an equivalence set since those predictions all stem from noisy copies of the same true theory). The thickness of the edges is a visual index of the product of the probabilities of the predictions connected by that edge. The test which can cut the sum of the thickness of multiple edges is chosen.

Suppose the test X_1 is chosen. It can distinguish the true (noiseless) predictions in left red EC and in lower green EC, which predict 0 in X_1, from the upper right blue EC which predicts 1. If the choice is actually 1, then all theories that predicted 0 are eliminated and their edges disappear, as shown in Figure 3.1b.

An important caveat: Note that when implemented, the EC^2 algorithm applied to the ECD will not

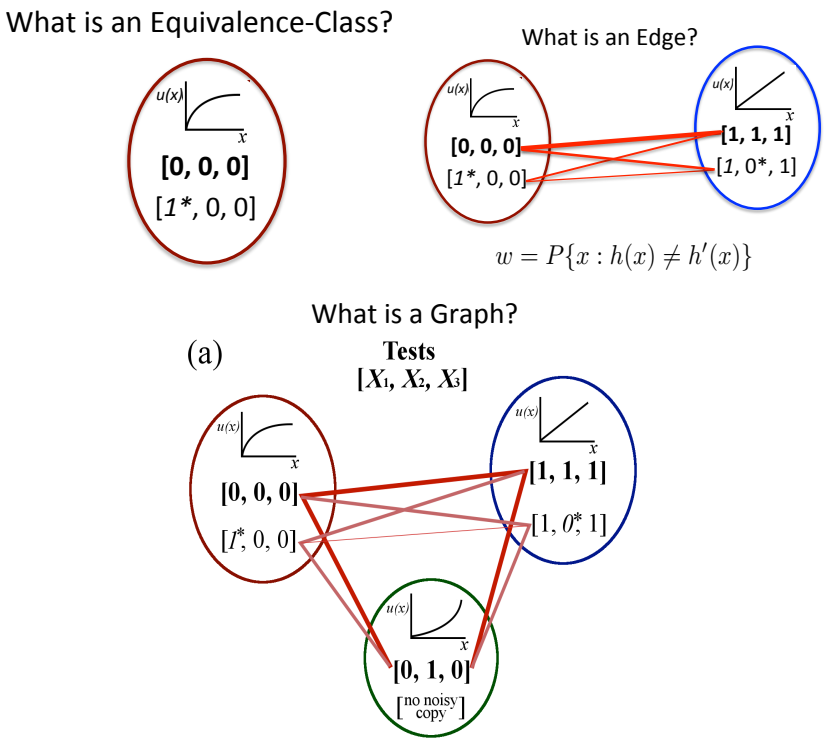


Figure 2.1: Constructing Equivalence Class Graphs. (top left) An equivalence class is the set of hypotheses that generate the same vector of outcomes. (top right) An edge joins two observation vectors generated by hypotheses from a sequence of tests in distinct equivalence classes. (bottom) An EC^2 graph consists of all equivalence classes and the edges connecting them.

always choose the true hypothesis. Here's an example illustrating why: Suppose there are 10 binary tests. A hypothesis h_1 is a vector of ten 0's and 1's. Now suppose we flip those numbers so that all possible vectors are generated (i.e., there are 10 one-flip noisy copies, 10!/(10-2)!2=45 two-flip noisy copies, etc.). If we do that for another hypothesis h_2, then one of the noisy h_1 copies will exactly match one of the h_2 copies. We cannot draw an edge between them, because no test can distinguish them. The approach can't guarantee that we will pick the true hypothesis even if we conduct all 10 tests.

Conceptually, our approach is therefore to flip some small percentage of all true hypothesis vector elements. This approach also can not guarantee perfect accuracy. The reason is that we might have failed to randomly flip a specific test choice for a true hypothesis in the way that corresponds to an actual noisy choice a person made.

The extent to which the wrong hypotheses are selected, averaged across many simulated designs, will be seen in ground-truth simulations below (Figures 2.4, 3.2, and 2.6). The asymptotic probability of choosing the correct hypothesis is about 90% for the reasonable noise rates we use. That is, even if all tests are conducted, the fact that noise is injected randomly means we sometimes choose a wrong hypothesis (if the bit flipping just happened to match choices). This is an unavoidable price for a realistic noise model such as this one, which

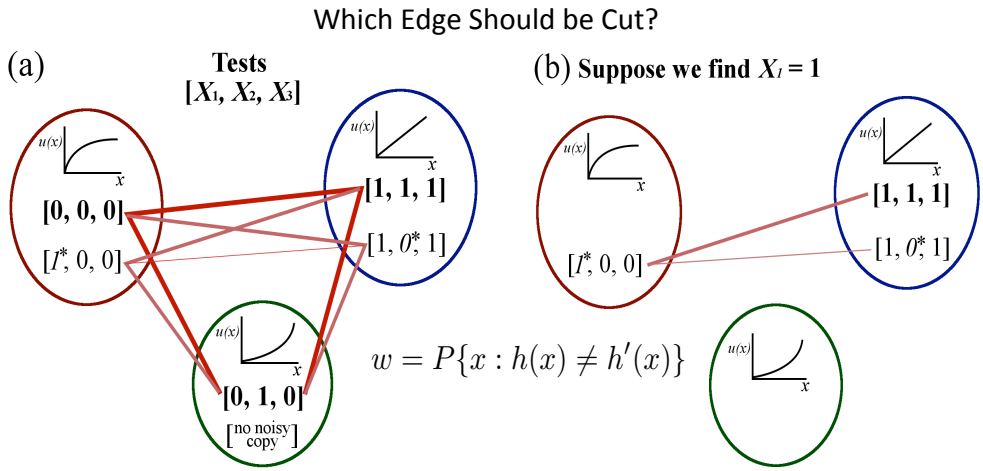


Figure 2.2: A stylized example of Equivalence Class Edge Cutting.

allows for human error to lead to model indistinguishability in some cases.

In practical implementation, for a reasonable number of hypotheses and observations, the equivalence classes can be exponentially large. In the next section, we'll discuss a couple of approaches to simplify the computations under mild assumptions.

2.4 Implementation: Efficiency and Challenges

The previous sections discussed the theoretical constructs by explicitly representing the observation vectors and edges between equivalence classes. In practice we can make simplifying assumptions, as discussed next, that leads to an effective implementation. In the following section, we show that by exploiting the adaptive submodular property of the objective function, we can implement the algorithm using lazy evaluation techniques that lead to orders of magnitude speedups.

2.4.1 Effective Implementation

For some noise models, Φ may have exponentially large support. In this case reducing Bayesian active learning with noise to Equivalence Class Determination results in instances with exponentially-large equivalence classes. This makes running EC^2 on them challenging, since explicitly keeping track of the equivalence classes is impractical. We next describe an algorithm that approximates EC^2.

For clarity, we only consider the 0-1 loss function, i.e. our goal is to find the most likely hypothesis (MAP estimate) given all data $x_{\mathcal{T}}$, namely $h^*(x_{\mathcal{T}}) = \arg\max_h P(h|x_{\mathcal{T}})$. Recall the defintion of the Bayesian noisy active learning objective and consider the weight between the edges between distinct equivalence classes $\mathcal{H}_i$ and $\mathcal{H}_j$:

$$w(\mathcal{H}_i \times \mathcal{H}_j) = \sum_{x_{\mathcal{T}} \in \mathcal{H}_i, x'_{\mathcal{T}} \in \mathcal{H}_j} P(x_{\mathcal{T}})P(x'_{\mathcal{T}}) = \big(\sum_{x_{\mathcal{T}} \in \mathcal{H}_i} P(x_{\mathcal{T}})\big)\big(\sum_{x'_{\mathcal{T}} \in \mathcal{H}_j} P(x'_{\mathcal{T}})\big) = P(x_{\mathcal{T}} \in \mathcal{H}_i)P(x'_{\mathcal{T}} \in \mathcal{H}_j) \tag{2.4.1}$$

In general, $P(X_{\mathcal{T}} \in \mathcal{H}_i)$ can be estimated to arbitrary accuracy using a rejection sampling approach with bounded sample complexity. We focus on the case where, upon observing all tests, the hypothesis is uniquely determined, i.e., $P(h_j|\mathcal{T})$ is deterministic for all $x_{\mathcal{T}}$ in the support of P. In this case, it holds that $P(X_{\mathcal{T}} \in \mathcal{H}_i) = P(\mathcal{H} = h_i)$. Thus, the total weight is

$$\sum_{i \neq j} w(\mathcal{H}_i \times \mathcal{H}_j) = \big(\sum_i P(h_i)\big)^2 - \sum_i P(h_i)^2 = 1 - \sum_i P(h_i)^2. \tag{2.4.2}$$

This insight motivates us to use the objective function

$$\Delta_{eff}(t|x_{\mathcal{A}}) = \big[\sum_x P(X_t = x|x_{\mathcal{A}}, \phi)\big(\sum_i P(h_i|x_{\mathcal{A}}, X_t = x, \phi)^2\big)\big] - \sum_i P(h_i|x_{\mathcal{A}}, \phi)^2 \tag{2.4.3}$$

which is the expected reduction in weight from the prior to the posterior distribution. Here $\phi \in \Phi$ is an instance of the noise model, thus $P(X_t = x|x_{\mathcal{A}}, \phi)$ is the likelihood that test X_t has an outcome x given

the observations $x_{\mathcal{A}}$ and noise ϕ. Note that the weight of a distribution $1 - \sum_i P(h_i)^2$ is a monotonically increasing function of the Renyi enropyt. Thus the objective Δ_{Eff} can be interpreted as a (non-standard) information gain in terms of the (exponentiated) Renyi entropy. In our experiments, we show that this criterion performs well in comparison to existing experimental design criteria, including the classical Shannon information gain. Computing $\Delta_{Eff}(t|x_A)$ requires us to perform one inference task for each outcome x of X_t, and $O(n)$ computations to calculate the weight for each outcome. We and present pseudocode of the greedy algorithm in Algorithm 1.

Input: Set of hypotheses $\mathcal{H}$; Set of tests $\mathcal{T}$; prior distribution P; function f.
begin
  $\mathcal{A} \leftarrow \emptyset$;
  while $\exists h \neq h' : P(h \mid \mathbf{x}_{\mathcal{A}}) > 0$ *and* $P(h' \mid \mathbf{x}_{\mathcal{A}}) > 0$ **do**
    foreach $t \in \mathcal{T}$ **do**
      $\Delta_{eff}(t|x_{\mathcal{A}}) = \big[\sum_x P(X_t = x|x_{\mathcal{A}})\big(\sum_i P(h_i|x_{\mathcal{A}}, X_t = x)^2\big)\big] - \sum_i P(h_i|x_{\mathcal{A}})^2$;
    Select $t^* \in \arg\max_t \Delta_{eff}(t|x_{\mathcal{A}})/c(t)$;
    Set $\mathcal{A} \leftarrow \mathcal{A} \cup \{t^*\}$;
    Observe outcome x_{t^*} and update posterior distribution over $\mathcal{H}$;

Algorithm 1: Sequential algorithm for effectively implementing EC^2.

Note that algorithm 1 is a framework for adaptive experimental design where starting with a prior of the hypotheses, we choose a test to run based on an objective criterion. After observing the outcome of the test, the posterior distribution over the hypotheses are updated. What differs here is that the objective criteria used is EC^2. Thus, any framework that is currently utilizing Information Gain can be easily replaced by EC^2 which has theoretical guarantees in the presence of noisy responses. Furthermore, as we will see in the next section, EC^2 objective function also lends itself to an accelerated greedy evaluations which makes it orders of magnitude faster than information gain.

2.4.2 Accelerated Greedy Evaluation

Implementation speeds of BROAD algorithms are important because subjects are waiting for dynamic computation after each trial. Therefore a relevant metric of comparison is the information gained as function of time (rather than just across the number of trials). We exploit the adaptive submodular property of EC^2 to develop an *accelerated* greedy implementation that reduces computation time.

The details of the algorithm are in Appendix B. The key insight (illustrated in figure 2.17, Appendix B) can be gleaned from the definition of adaptive submodularity: $\Delta(t \,|\, \mathbf{x}_{\mathcal{B}}) \leq \Delta(t \,|\, \mathbf{x}_{\mathcal{A}})$, where test $t \in \mathcal{T}$ for all $\mathbf{x}_{\mathcal{A}} \prec \mathbf{x}_{\mathcal{B}}$. The value of a test t can only decrease in future rounds as more observations are made. Initially, the values of all the tests are calculated and sorted[21], and the highest value test is presented in the first round

[21] A *Priority Queue* is often used to store sorted values for efficient calculation.

(and removed from the queue). In subsequent iterations, the value of the most valuable test is recalculated, and only the tests that had a higher value than this in the previous iteration needs to be recalculated. This saves an enormous amount of computation, and in practice could lead to orders of magnitude improvement in efficiency.

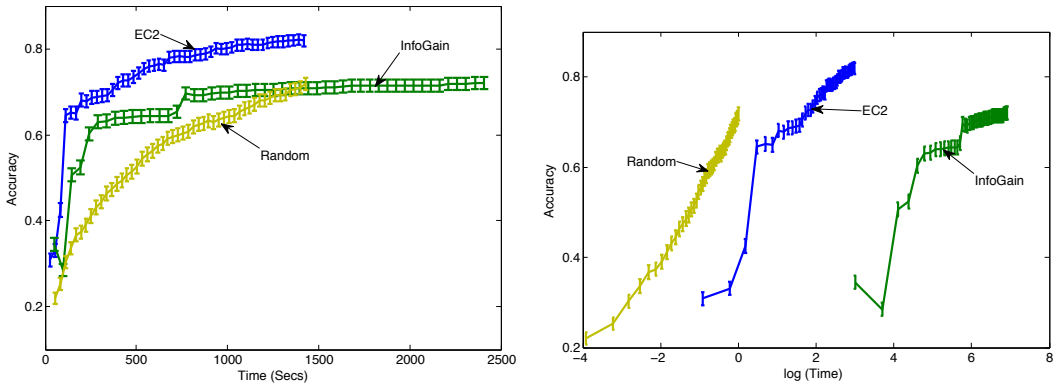


Figure 2.3: (Left) Accuracy as a function of time assuming a response time of 28 seconds (average response time across all subjects). (Right) Accuracy as a function of logarithmic time with no response delay.

In figure 2.3, we plot the Information gained per unit of time for accelerated EC^2, Information Gain and Random, assuming a response time c (= 28 seconds) per question. EC^2 is more efficient compared to Information Gain and Random. When the responses are immediate ($c = 0$), such as when querying a machine database, Random performs well in a short time assuming no costs for the number of queries asked. EC^2 is superior to InfoGain in both time and accuracy in this case.

Method	Theoretical Properties	Reference
Information Gain	Equivalent to GBS in noise free setting.	Lindley (1956)
Value of Information	Can only be optimized in restricted settings	Howard (1966)
Generalized Binary Search (GBS)	Submodular in non-adaptive setting assuming conditionally independent observations	Burnashev and Zigangirov (1974)
Random	Asymptotic convergence when all tests applied	–

Table 2.2: Summary of the experimental design methods that are compared against EC2

2.5 Design and Evaluation

The goal is to adaptively select a sequence of tests to present to a subject, in order to distinguish which of the six theories best explains the subject's responses. Each choice test t is a pair of lotteries, (L_1^t, L_2^t). Based on the theory that represents behaviour, one of the lotteries would be preferred to the other, denoted by a binary response $x_t \in \{1, 2\}$. (Stochastic response will be introduced shortly.)

Each lottery consists of three outcomes and their corresponding probabilities, $(L_i^t) = \{(x_1, p_1), (x_2, p_2), (x_3, p_3)\}$. The middle outcome $x_2 = \$5$ is fixed in our experiments. The loss outcome x_1 varied between $\{-20, -10, -5\}$ dollars and the gain outcome x_3 varied between $\{10, 20, 50\}$ dollars. The distribution (p_1, p_2, p_3) over the payoffs was varied, where the probabilities $p_i \in \{0.01, 0.1, 0.25, 0.5, 0.75, 0.9, 0.99\}$ are chosen such that $p_1 + p_2 + p_3 = 1$.

To conduct the experiment, we used a display that showed the two lotteries along with a bar graph (see Appendix A). By considering all non-identical pairs of such lotteries, we obtained the set of 31626 possible choice test (i.e. question pairs).

2.5.1 Ground Truth Analyses

In this section we evaluate the quality of five different design methods, which are summarized in Table 2.2. Both Information Gain and Generalized Binary Search are adaptive submodular in the deterministic no-noise setting (where running a test long enough is guaranteed to generate the correct response), Golovin and Krause (2010b). In the Random method, a test was picked uniformly at random in each trial from the set of possible tests without replacement. This should asymptotically converge to the optimal decision once all tests have been conducted. Value-of-Information (VoI) criteria is defined w.r.t a decision problem to choose which among several, potentially expensive, variables to observe in order to most effectively increase the expected utility (Howard (1966)). Although maximizing VoI non-myopically is optimal, in all but chain-models (e.g. Markov Chains, and slight generalizations) this non-myopic optimization is extremely intractable (Krause and Guestrin (2009)), making it infeasible for our problem.

We evaluated the speed and ability of the five methods to recover the true model based on a large sample of

simulated response sequences. In each sequence, a specific true model is assumed, which generates simulated responses to the sequence of test choices generated by each method. The conclusion the method draws about which model is likely to have produced the observed responses is then compared to the true (known) model. This model-recovery exercise is sometimes called "ground truth" analysis. The responses were generated using a trembling hand function, with a known probability ϵ of making the wrong choice due to noise.

We chose discrete prior parameter values for different theories based generally on previous estimates reported in empirical literature, as follows:

∘ The expected value (EV) model has no parameters. ∘ CRRA has one risk-aversion parameter, $\Theta_{CRRA} \in \{0.2, 0.4, 0.6, 0.8, 1.0\}$. ∘ A range of parameter values for prospect theory has been reported in the literature (e.g., Booij *et al.* (2010), Table 1). For both PT and CPT we chose $\rho \in \{0.5, 0.62, 0.74, 0.86, 0.98, 1.1\}$, loss aversion, $\lambda \in \{1, 1.5, 2, 2.5, 3.0\}$, and Prelec-style probability weighting inflection $\alpha \in \{0.4, 0.52, 0.64, 0.76, 0.88, 1.0\}$. ∘ The range of parameter values for moments models MVS and SMVS were not readily available in experimental literature. To choose an appropriate range of parameter values, we created artificial choice data based on prospect theory parameters (from the last paragraph) and found MLE best-fitting MVS and SMVS parameters.[22] The resulting parameter ranges were: $w_\mu \in \{0.6, 0.7, 0.8, 0.9, 1.0\}$, $w_{\sigma^2} \in \{0.05, 0.25, 0.45, 0.65, 0.85, 1.05\} \times 10^{-2}$ and $w_\nu \in \{0, 1, 2, 3, 4, 5\} \times 10^{-4}$. For SMVS the parameter ranges chosen were: $w_\mu \in \{0.6, 0.7, 0.8, 0.9, 1.0\}$, $w_\sigma \in \{0.05, 0.15, 0.25, 0.35, 0.45\}$ and $w_{\tilde{\nu}} \in \{0.0, 0.1, 0.2, 0.3, 0.4\}$. We added a slight offset to w_{σ^2} and w_σ of 0.05 and 0.0005 respectively, to distinguish the model predictions between distinct parameter values (otherwise several values would make the identical prediction over all designs). Note that this procedure actually limits the ability of the BROAD model to find questions that sharply distinguish PT and moments models, because the candidate parameters for the moments models are those which fit PT-generated choices. However, once more information about likely moments models parameters is built up (from our data and others) the entire procedure can be repeated with a different parameter range that should be even more diagnostic than the one used here.

For the ground truth analysis, each of the six model classes started with an equal overall prior belief of $1/6$, spread uniformly across all parameter value configurations in that model class. (For example, each of the 125 specific parameter configurations for PT was assigned prior $1/6$ divided by 125.) In each ground truth run, a model and parameter configuration was chosen at random. Thirty test choices were then generated using each of the five methods. In each such run, we record whether each of the five methods is accurate after each

[22]For all possible experimental designs (roughly 32,000) in our setup, we generated the responses for all possible parameter values (125) for the prospect theory model, which generated about 4 million artificial datapoints. Next, we used a binary search to find a range of good-fitting parameter values for the moments models. We started with a wide range, and estimated the maximum likelihood fit of the generated responses. Then we did a binary split, by choosing the half of the range that contained the MLE. We repeated this procedure until dividing the parameter range led to a decrease in the likelihood of the responses, i.e. some responses are better explained with the higher parameter range.

question. Accuracy in a run is defined as assigning the highest posterior probability to the ground-truth model (for the overall model in the same class) in that run. The procedure was repeated for 1000 runs.

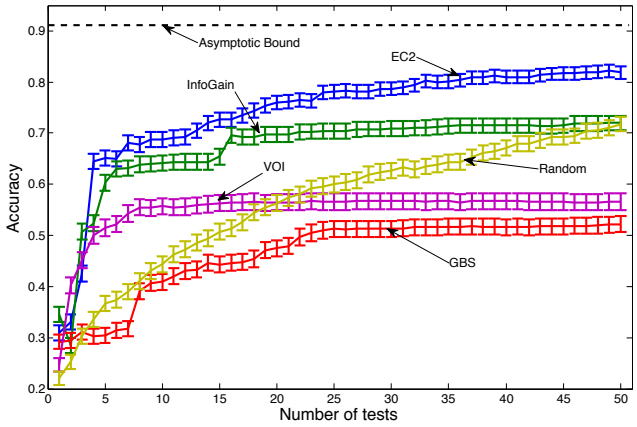


Figure 2.4: Performance of the algorithms when probability of error, $\epsilon = 0.05$. The asymptotic bound from running all tests in random order is also shown.

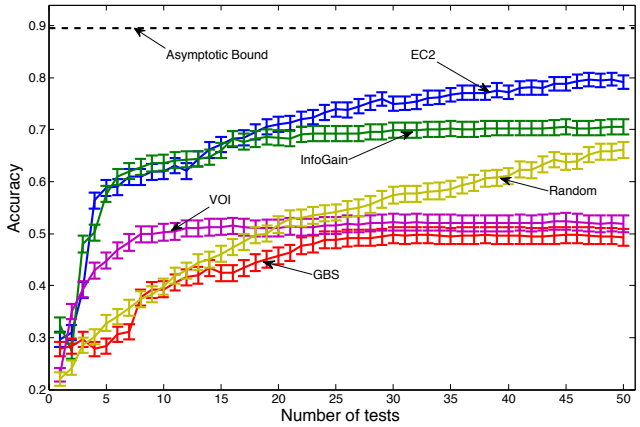


Figure 2.5: Performance of the algorithms when probability of error, $\epsilon = 0.10$. The asymptotic bound from running all tests in random order is also shown.

Figures 2.4, 3.2, and 2.6 show the performance of the 5 methods for three error rates ($\epsilon = 0.05, 0.10, 0.20$).[23] We also plot the asymptotic performance upper bounds from running all the test choices (they are 91.09%, 89.9% and 88.2% across the three error rates).

The Figures show that VoI and GBS both perform significantly worse than Random in the presence of noise. These methods are clearly poorly-suited when responses are noisy, which most experimental social science data are likely to be. Indeed, the accuracy of the random method degrades as noise increases. EC^2 is noticeably more accurate than InfoGain for lower error rates, and comparable for the highest error rate

[23] Test-retest switching rates for monetary gamble choices repeated twice are similar to these error rates, around 15% (e.g. Camerer, J Risk Uncertainty 1989).

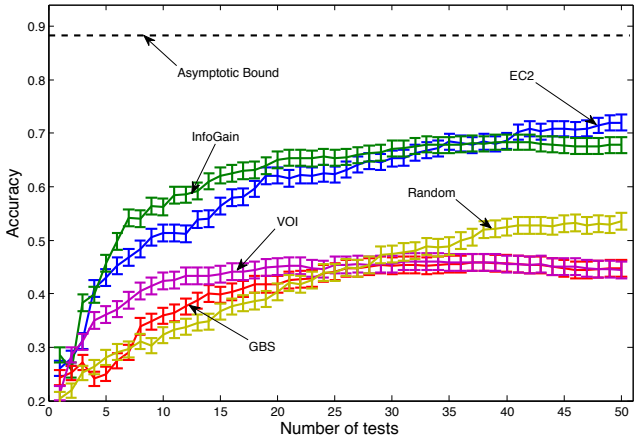


Figure 2.6: Performance of the algorithms when probability of error, $\epsilon = 0.20$. The asymptotic bound from running all tests in random order is also shown.

$\epsilon = 0.20$. It is also evident that most of the improvement in accuracy occurs in the first 5-20 test choices; and improvement is most rapid when error is low.

These ground-truth analyses show how frequently the test selection methods get the right answer. Another performance criterion is how well the experimental setup and methods distinguish among the models. To investigate this, we plot the accuracy curves for the individual models from the ground truth process in Figure 2.7, for $\epsilon = 0.05$. The experimental setup we use is extremely efficient (almost 100% accuracy) at classifying SMVS, EV and CRRA. However, classification accuracy for CPT is weaker and for MVS classification is relatively poor, although significantly better than chance (which is 1/6 $\approx$ 16.7%).

–	PT	CPT	MVS	NMVS	EV	CRRA
PT	72.7	21.5	2.9	0	2.9	0
CPT	40.6	57.3	1.1	0	0.5	0.5
MVS	0.6	0	52.7	5.8	40.3	0.6
NMVS	0	0.6	1.3	96.2	0.6	1.3
EV	0.5	0	2.6	0.5	96.4	0
CRRA	0	0	0	0	0	100.0

Table 2.3: Misclassification Matrix: Percentage of samples (row) that are classified as another model (column) using EC2 after 50 rounds. The diagonal represents correct classification.

We analyzed the classifications to generate a *misclassification matrix* in Table 2.3, i.e. we looked at how the samples from the models are classified by EC^2. Each matrix entry shows the frequency with which a true row model is classified into each possible column model (Excellent classification would be evidenced by percentages close to 100% on the diagonal.). As pictured in Figure 2.7, the method correctly classifies SMVS, EV and CRRA almost perfectly. The analysis reveals where the misclassification occurs. One limit is that a significant number of PT samples (21.5%) are misclassified as CPT, and 40.6% of CPT samples are

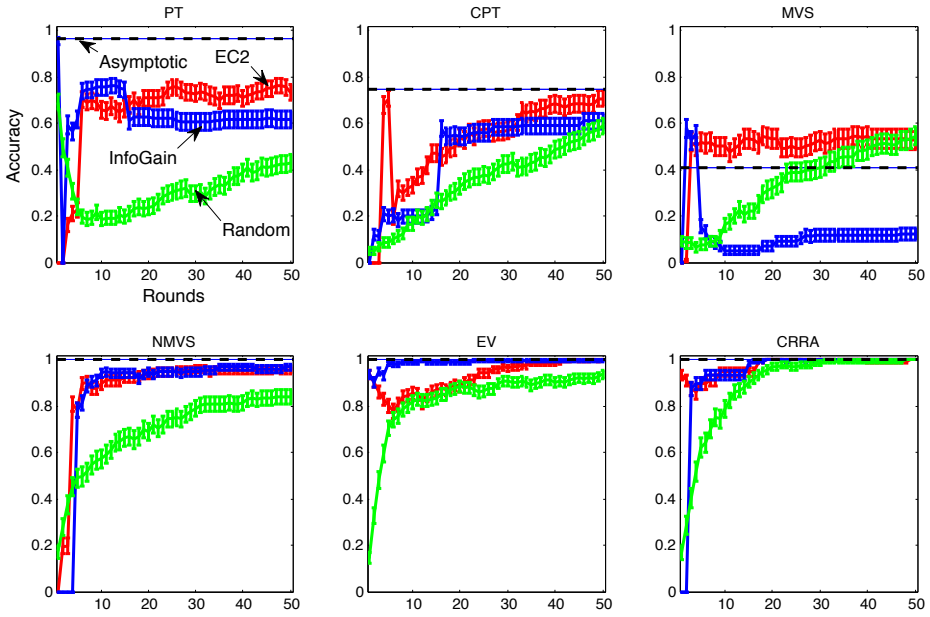


Figure 2.7: Performance of the algorithms for individual models when probability of error, $\epsilon = 0.05$. The asymptotic bounds are also shown for recovering individual models.

misclassified as PT. The other limit is that a large number of MVS samples (40.3%) are misclassified as EV. This is because the weight on skewness in MVS can be zero, and when the weight on variance is low enough, the choices of EV and a low-variance-weight MVS model are difficult to distinguish.

The efficiency of classification can potentially be improved through alternate experimental setups. Classification quality will also be sensitive to how the parameter ranges are chosen, to how reliably subjects make choices consistent with one model or another, and other factors. Finally, keep in mind that even with an efficient setup, the method cannot necessarily achieve perfect classification using just a small number of test choices.

2.6 Manipulation

In BROAD designs, subjects could misrepresent their true preferences in some choices *if* the following chain of conditions hold: (i) they believe that future test choices depend on previous responses; (ii) they can compute how to misrepresent preferences in early choices to create better future choices (as evaluated by their own preferences); and (iii) the value of misrepresentation is high enough to be worthwhile.

We present arguments and evidence that misrepresentation resulting from the chain of conditions (i)-(iii) is likely to be rare. When misrepresentation does occur, it could be easily detected and is not likely to lead to wrong conclusions about revealed preferences which cannot be easily reversed.

We illustrate with a concrete example from our empirical study. For maximum fidelity, the reader should assume she has true preferences corresponding to prospect theory (PT), that was the most common type in the data.

The first test choice presented to all subjects is the same (because no choice history is used to fine-tune the test-choice selection). Subjects are asked to choose between $L_1 = \{(-10, 0.9), (5, 0), (10, 0.1)\}$ and $L_2 = \{(-20, 0.5), (5, 0.49), (50, 0.01)\}$.

What would a myopic (non-manipulating) agent choose? What would a manipulating agent choose? In post-experiment polls of this type with subjects people are not clear how manipulation would change the answer about what risky choice to pick. Even if the algorithm is known, solving the decision tree to make the optimal strategic decision is not easily computable. Here's an easier version: Suppose your preferences are EV. The maximum EV choice is L_2. Should you strategically choose L_1 instead?

These examples might shake the readers' faith that manipulation is easy to figure out. But more can be said. Next we speculate about whether manipulation pays, whether it occurs, how it might be detected in experimental results, and how its effects might be muted or corrected.

◦ Does manipulation pay? To partially answer this question, it is helpful to establish an upper bound on the maximum gain from manipulation, for a particular design and player type. The upper bound on the marginal gain is likely to be low. Here's why: In later periods, it does not pay to manipulate since doing makes suboptimal immediate choices. And in early periods, manipulation is immediately costly for the same reason. So there is a natural tradeoff between the cost of manipulation in a period– the utility losses from deliberately making the wrong choices– and the future gains from improved choice sets. It could be that in a 10-period experiment, for example, manipulation is only beneficial in the first three periods. If so, the posterior probabilities computed after 10 periods might be close to the correct posteriors because they include 7 periods of non-manipulation choice data after three periods of misleading data. It is also possible that when ranking different subjects by their risk-aversion, for example, we can recover an approximately correct ranking across

subjects even if manipulation leads to biased estimates of their means.

To be more concrete, we computed the upper bound of manipulation for a *clairvoyant* agent. Specifically, consider a version of our experimental setup with 10 questions, leading to $2^{10}(= 1024)$ potential sequences of choices. We assume that the agent knows all the conditional sequences of tests, or tree, that will be generated by the EC^2 algorithm. He chooses the sequence of responses that maximize his expected earnings over all 10 trials.

We find that the earnings of a clairvoyant agent with EV preferences are only 8% higher than an agent who myopically chooses options that maximize his earning for each round. More importantly, the resulting type classification is incorrect in 68 cases, or with 6.6% error rate, which would partly be due to sampling error (Table 2.3).

◦ Does manipulation actually happen? One way to see whether manipulation is common is to do an experiment in which one treatment group is told that an adaptive algorithm is used to generate their future choices, and it may benefit them to choose risks they do not like to induce the algorithm to offer them better choices in the future. We created two treatments with different instructions. Subjects were told that the algorithm is adaptive and can theoretically be manipulated (Additional instructions in Appendix A). In post experiment self-reports, most subjects reported that they did not try to manipulate the system or were not able to do so. Others said they did make different choices (See online supplementary methods for self-reports). The most striking feature of the result was that the EC^2 algorithm identified more subjects in the "informed" as using expected value maximization. Some subjects said that knowing an algorithm was changing their choices led them to maximize EV. It is also notable that the subjects in the "informed" condition did not have higher average earnings (as might be expected if they are manipulating). This is explained in more detail below.

◦ Can manipulation be detected? Strategizing will typically leave clear fingerprints in the data from choices across a sequence of questions. First, in typical cases the posterior probability of the most likely model (as judged from final round result) goes up across the trials. In contrast, a manipulating respondent will appear to be one "type" in early trials, and then revert to their true "type" in later trials (as the future benefit of manipulation shrinks). This will leave a telltale pattern of posterior probabilities veering from one type to another, from earlier to later trials. We explored how diagnostic this pattern is for simulated agents who optimally misrepresent, and in actual experimental data, in figure 2.14. When artificial subjects are known to be manipulating, the percentage of subject classifications corresponding to the true type (as judged from final-round posteriors) is very low in early rounds. However, the opposite pattern is evident in the actual choice data.

Furthermore, if manipulation in early choices is cognitively demanding, manipulating subjects will initially

have high reaction times that decrease sharply over time, as they begin to express easily-computed true preferences in alter trials. However, reaction times from actual experimental data are similar in the control and cued-manipulation treatment, see Figure 2.15.

◦ How can manipulation be limited? There are two sensible ideas about how manipulation could be limited. (Unfortunately, both of them weaken the incentives to respond truthfully to all questions that are asked.) The first remedy[24] is to create a space of questions Q, then sample from Q using a BROAD procedure. Subjects are told that one question will be chosen for payment *from the entire set Q*, not from the subset that were actually selected and answered. Misrepresenting preference in the first few questions in the BROAD procedure *does* lead to "better" questions being included later in the BROAD set, but it does not change at all the chance that those questions will be chosen for payment (since BROAD-selected and all other questions in Q are equally likely).

A second remedy is to make the probability with which questions are chosen for payment much higher for the earlier questions in the sequence (where the incentive for mispresentation is highest). We will investigate numerically how much "front-loading" probability onto early questions mutes the incentive to misrepresent.

To tackle this, we ran 2 sessions, with 30 subjects at UCLA, with augmented instructions (Appendix A). Subjects were informed that an adaptive algorithm was being used, and that prior choices will affect subsequent options.

By definition, subjects would manipulate the algorithm in order to increase their earnings. The average earnings in the prior session was 9.44 dollars, with a standard deviation of 9.40, and the average earnings with the augmented instructions was 9.00 dollars with a standard deviation of 9.3. Thus the earnings from both sessions were nearly identical.

In Figure 2.8 we show the classifications for the two different cases to test for effects of manipulability (see Section 2.6 below). We cannot reject the hypothesis that the two distributions of classifications are statistically different using a 2-sided Kolmogorov-Smirnov test. We show stacked plots of the posterior distribution of the two treatments in figure 2.9.

Furthermore, we obtained self-reports from subjects (the questions are shown in Appendix A), to find out if the participants used any strategies to game the system. The majority of the participants did not try to manipulate the algorithm. The few subjects who reported attempts to manipulate the algorithm did not come up with any effective method (summarized self-reports to be made available online).

[24]This idea was suggested by Cathleen Johnson, but never published.

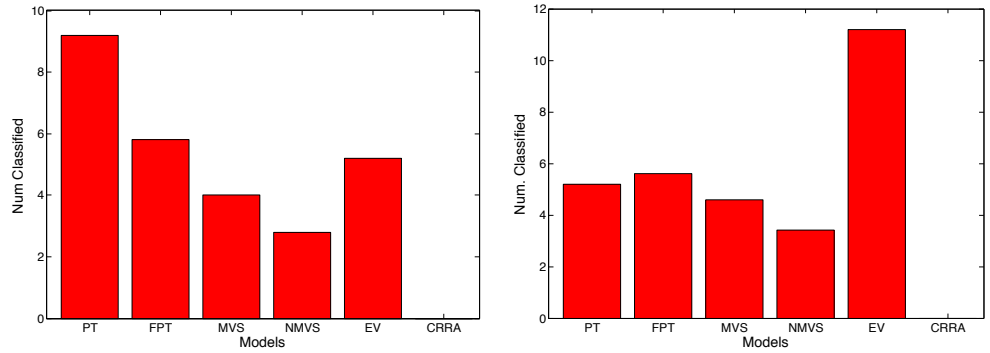


Figure 2.8: The classifications for the two treatments to test for manipulability. Treatment 1 has 27 subjects, and treatment 2 has 30 subjects. The difference between the two distributions was not found to be statistically significant.

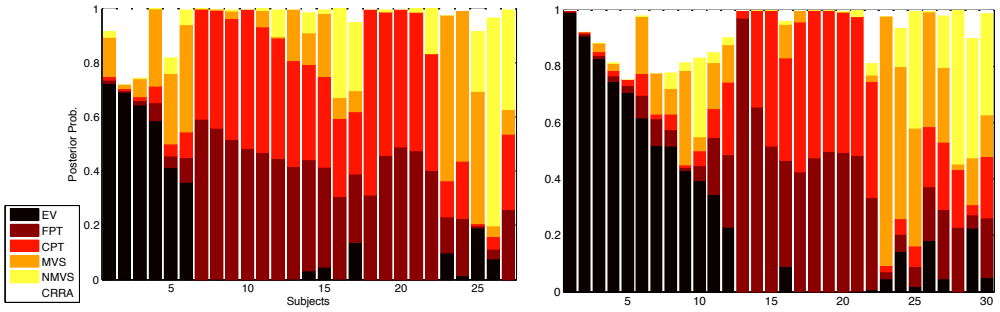


Figure 2.9: Stacked bar plot of posterior distribution over the 6 models (averaged over the final 5 rounds) for 27 subjects in the non-manipulation condition (left) and the manipulation condition (right).

2.6.1 Noise Specification

The theoretical guarantees for EC^2 do not depend on the choice of the noise model Φ, but the specific objective criteria implemented in section 2.3 assumes a conditionally independent noise model. In the ground truth analysis above, the errors are generated with a probability of ϵ, and responses are correct with probability $1 - \epsilon$. A natural concern is, does the ϵ noise model accurately capture the noisy responses of subjects, and if not, does it introduce a bias in the results?

A popular alternative choice model is *softmax* (for e.g. used in computational neuroscience by Sutton and Barto (1998)) where the probability of chosing an option c is $P(c) = \exp(U(c)/\sigma)/\sum_i \exp(U(c_i)/\sigma)$. As σ grows, the subject will be less likely to choose the most utile alternative and make an error. If two choices have similar utility, then the subject is more likely to make errors, which can partially be explained by the fact that the subject is indifferent between the two outcomes with similar utility. We test whether the resulting classifications are robust when the noise generation process is softmax, and where the misclassifications (if any) occur. To do this, we run the ground truth analysis again, with the noise now being generated by a softmax

function with parameters, $\sigma = \{0.86, 1.12\}$[25].

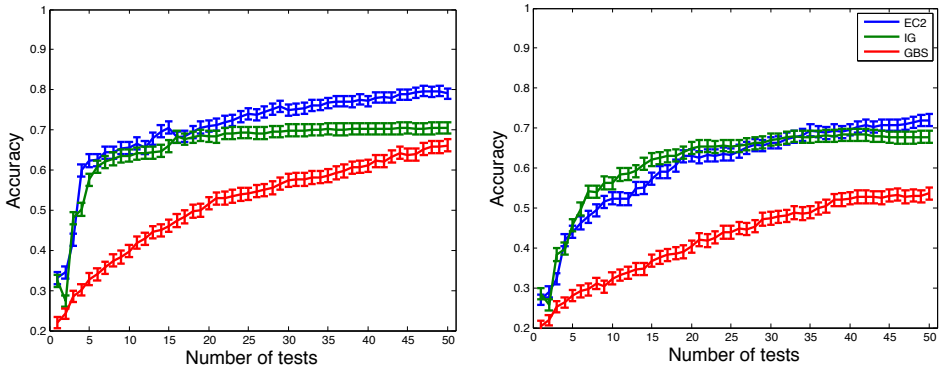


Figure 2.10: Performance of the algorithms for individual models when probability of error, $\sigma = 0.86$ for an average error rate of 10%(left) and $\sigma = 1.12$ for an average error rate of 20% (right).

–	PT	CPT	MVS	NMVS	EV	CRRA
PT	-2.1	-2.5	-0.9	0	0.9	0
CPT	-9.6	-6.3	0.2	0	0.0	0.0
MVS	0.6	0	-2.8	-0.8	-7.3	0.6
NMVS	0	0.0	-0.3	-3.0	0.0	-0.3
EV	0.1	0	1.1	0.1	1.1	0
CRRA	0	0	0	0	0	0

Table 2.4: Differences in Misclassification Matrix: Percentage of samples (row) that are classified as another model (column) using EC2 after 50 rounds for ϵ noise minus the Softmax noise model for average error rate of 0.10. The diagonal represents no difference in classification accuracy.

Figure shows the accuracy of EC^2, Information Gain, and random questions with the number of rounds for $\sigma = \{0.86, 1.12\}$ corresponding to average error rates of 10% and 20%. The difference in accuracy when the noise model is ϵ or softmax is statistically not significant. We compare the misclassification matrix from the softmax noise model with the ϵ noise model, and the differences are shown in table 2.4. Each cell is the difference in misclassification by the ϵ noise model, and the softmax noise model where average error rate is 0.10. The fact that there is no overall difference in accuracy is partially understandable since we picked values of σ so that the average error rates match the ϵ error rates. We notice fewer classification errors with the softmax noise model since the EC^2 criteria does not choose tests, at least in the initial rounds, where the utility of the choices are very similar. The ϵ noise thus "throws off" the algorithm more since choices with high difference in utility can still generate errors (at random) whereas they are unlikely in the softmax model.

[25]We simulated choices over all tests and all theory and model parameter values for values of σ between 0.5 and 2 in increments of 0.01. We calibrated the σ such that the average number of errors over 1000 trials for $\sigma = \{0.86, 1.12\}$ are 0.10, and 0.20 respectively.

2.7 Empirical Data and Analysis

We ran experiments with 57 subjects in 4 separate sessions at UCLA to determine which models and parameters fit their behaviour best. We used 2 different treatments using slightly different instructions (shown in Appendix A).[26]

Thirty test choices were created using EC^2. To incentivize the subjects, one of their choices was selected at random for payment. Subjects were given an initial endowment of $20 and a $10 show-up fee[27]. Subjects received payment based on the lottery they picked in the selected choice, and on a random outcome from that lottery[28]. Earnings from choices ranged from $5 to $50, with a mean of $9.21 and median of $5.

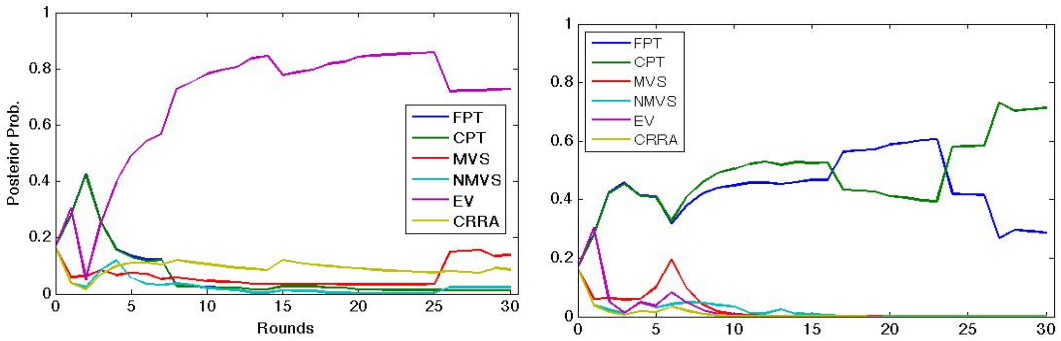


Figure 2.11: Examples of the dynamically-estimated posterior probability over all models for two subjects. The subject on the left is rapidly classified as EV; the one on the right is less rapidly classified as CPT.

The data reveals heterogeneity in the subjects' estimated types, and the dynamics of how type estimation changes over the 30 test choices. We illustrate two examples in Figure 2.11. In one example the subject's type is identified very early in about 10 rounds. In the second case, the classification is evenly split between PT and CPT until the final five rounds.

Figure 2.12 shows stacked plots of the posterior distributions of all 57 subjects across all six models (averaged over the last five experimental trials). The stacked plot reveals high confidence in the EV classifications. The posterior mass is typically spread between the PT and CPT classifications, for those cases in which either of those models are considered the most likely. The moments models MVS and SMVS fit fewer subjects, but are typically sharply distinguished from each other. It is perhaps surprising that CRRA is never the best-fitting model.

Overall classification can be summarized by compressing the data in Figure 2.12. Figure 3.4 shows

[26] We followed the IRB protocols of UCLA for the conduct of human subject experiments.

[27] The earnings from the experiment were added to the endowment. Since the minimum outcome was minus $20, the subjects would never have to pay money, as per IRB requirements.

[28] A 10×10 matrix of values from 1 to 100 was generated by random permutation, and subjects were asked to pick a row and column number. The number at the row and column was the random probability for the subject, which was used to realize the lottery outcome. For example, if the lottery picked was $L = \{(-5, 0.1), (5, 0.5), (20, 0.4)\}$, and the random number generated was 67, the subject would earn $20.

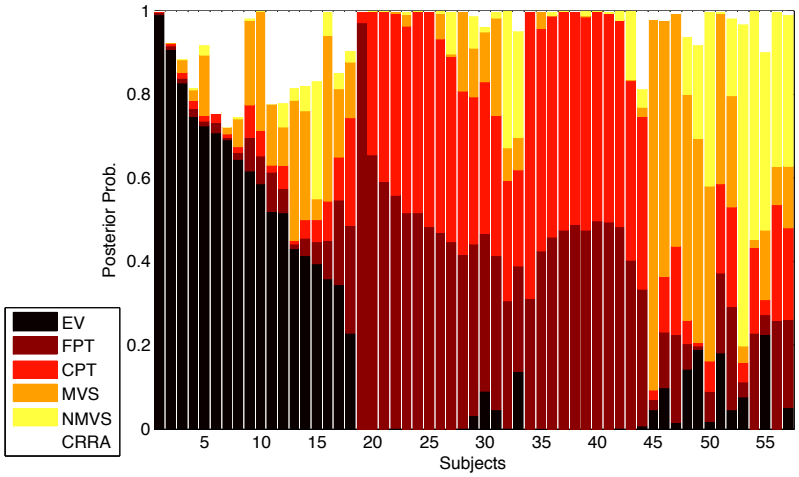


Figure 2.12: Stacked bar plot of posterior distribution over the 6 models (averaged over the final 5 rounds) for 57 subjects. The subjects are ordered on the x-axis based on maximum posterior probability into models EV, FPT, CPT, MVS, SMVS and CRRA .

the average posterior distribution of all models after each round, aggregated over all subjects. It is clear that the posterior distribution of EV, PT and CPT are generally higher than the moments models and CRRA. Furthermore, it appears that separation in model fit both occurs rapidly, in about five trials, and stabilizes after about 20 trials; after 20 trials the overall ranking of the model fits does not change at all.

2.7.1 Distribution of Parameters

At the end of the experiment, we obtain a posterior distribution over all hypotheses (theory and specific parameters). We sum over all parameters within a theory to obtain the posterior distribution over the theories, and obtain the MAP classification of the subjects. Given the final classification to a theory, we can also obtain the MAP estimate of specific parameters. Expected Value (EV) has no parameters and $n = 17$ classifications, and Constant Relative Risk Aversion (CRRA) had no classifications. For the rest of the theories, the MAP parameter estimates were:

- Functional Prospect Theory (PT) had $n = 14$ classifications; average values of parameters (with standard errors) were: $\rho = 0.88(\pm 0.07), \lambda = 2.22(\pm 0.40), \alpha = 0.77(\pm 0.12)$.
- Cumulative Prospect Theory (CPT) had $n = 11$ classifications; average values of parameters (with standard errors) were: $\rho = 0.79(\pm 0.11), \lambda = 1.90(\pm 0.39), \alpha = 0.85(\pm 0.09)$.
- Unnormalized Mean-Variance-Skewness (MVS) had $n = 9$ classifications; average values of parameters (with standard errors) were: $w_\mu = 0.89(\pm 0.10), w_{\sigma^2} = 0.0026(\pm 0.0023), w_\nu = 0.000028(\pm 0.00001)$.

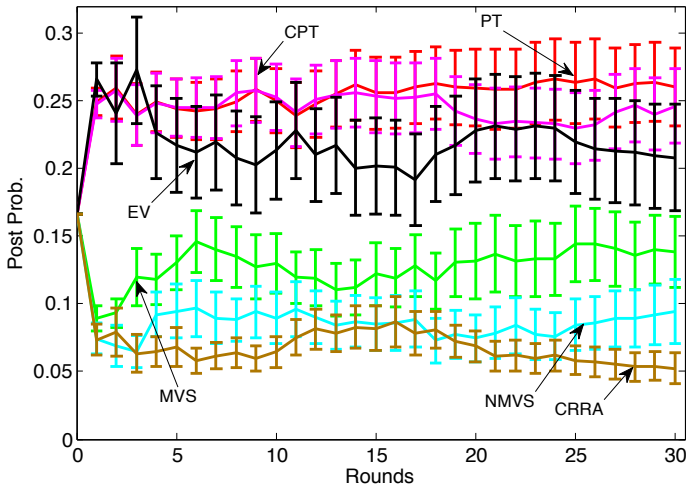


Figure 2.13: The posterior distribution of the 6 models after each round. The aggregate mean and standard error is shown for 57 subjects.

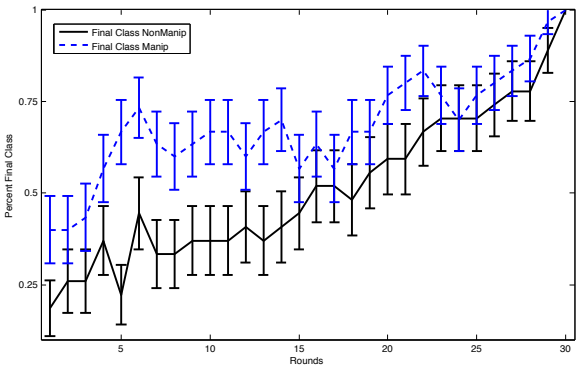


Figure 2.14: The percentage of subjects whose current classification in a round, corresponds to final classification (in round 30). Subjects trying to manipulate the algorithm might misrepresent their true type in early rounds.

- Normalized Mean-Variance-Skewness (NMVS) had $n = 6$ classifications; average values of parameters (with standard errors) were: $w_\mu = 0.94(\pm 0.18), \sigma = 0.188(\pm 0.08), w_{\tilde{\tau}} = 0.16(\pm 0.07)$.

Since the objective for EC^2 is to distinguish between theories, the experimental design and the sequence of questions asked is not optimized towards increasing the resolution of the parameter estimates, or even narrowing in on the correct parameters. Our methodology can easily be extended to estimating the parameters of a specific theory, such as prospect theory, with a high resolution.

2.7.2 Interpretation of Results

How should the experimental results be interpreted in the context of results in literature? Our lab experiments with 57 subjects found no classifications, and least evidence, for Constant Relative Risk Aversion (CRRA).

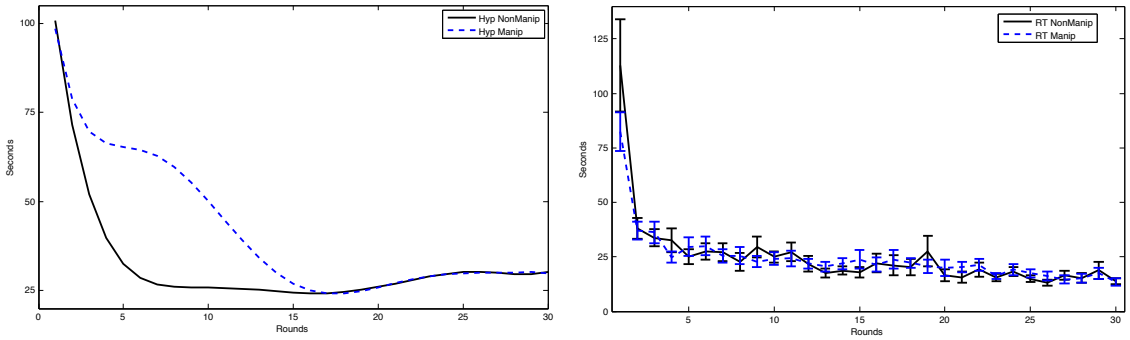


Figure 2.15: (left) A cartoon of the hypothetical reaction times in the two treatments to test for manipulation. We expect higher reaction times when subjects are trying to manipulate. (right) The actual reaction times with standard error bars for the two treatments.

This is a surprising finding in the context of the literature where CRRA is the most popular model used in lab experiments, including in the HL paradigm (Holt and Laury (2005)). A plausible explanation is that CRRA is used to estimate a subject's risk aversion in risky choice experiments that don't include losses or outcomes with very low probabilities. Our experiments suggest that since Prospect Theory nests CRRA, these subjects are likely Prospect Theory types that exhibit risk aversion when the choices are only in the domain of gains without extreme probabilities.

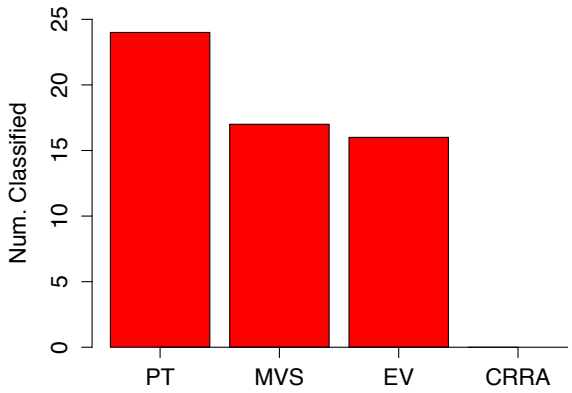


Figure 2.16: The "collapsed" classification of 57 subjects into the four class of theories: Prospect Theory (PT), Moments Models (MVS), Expected Value (EV) and Constant Relative Risk Aversion (CRRA).

The misclassification matrix in table 2.3, shows that with a small number of rounds (30, in our experiments), the two variants of prospect theory can be hard to distinguish, and the two variants of moments models are also hard to separate. Since we obtain the posterior distribution over the theory and parameter values, we can consider "collapsing" the classifications into 4 distinct class of theories: Expected Value (EV), Prospect Theory (PT), Moments (MVS) and CRRA. Doing so is as simple as summing over the hypotheses over two theories equivalent to adding posterior probabilities. Figure 2.16 shows the "collapsed" classification according to the

4 classes of theories.

We find that PT is the dominant class of theories, although there is good evidence for EV and Moments models. We have to be cautious in making strong claims about which theory is better at classifying subjects since the classification is mixed (and there genuinely are mixed types in the population), and because the results depend on the experimental designs. In our designs, there are 3 possible outcomes (states) and the probabilities are varied over them, so we can compute the first, second and third moments. dAcremont and Bossaerts (2008) has shown that as the number of states get very high, subjects switch to computing statistical moments rather than tracking states (thus state representations collapse to moments). This is relevant in finance since the number of states is very high, and calculating utility using prospect theory is computationally intensive.

The heterogeneity in classifications in our experiments is consistent with other studies that have found heterogeneity in risk and loss aversion in larger populations; Some of the heterogeneity can be explained by socioeconomic factors (Von Gaudecker *et al.* (2011)). Figure 3.4 shows the posterior probability for the 6 theories aggregated for all 57 subjects, which can be interpreted as our confidence level in the classifications. There is a possibility we have not considered a theory that would be an accurate model of the choices for most subjects. In that case, the "unexplained" evidence would be captured by the noise term.

Finally, our conclusions are based on the assumption that the subject's strategy (or type) stays consistent throughout the experiment. If the subject's strategy changes arbitrarily throughout the experiment, then it is generally difficult to classify them. However, if the subject's strategy changes early on and then stays consistent till the end of the experiment, since we utilize Bayesian updates, and account for noise (so a hypothesis does not get eliminated entirely, but the evidence is decreased), the algorithm can recover to find the subject's final strategy.

2.8 Conclusions and future directions

This paper proposes and implements a new BROAD method for adaptively selecting which choice pairs to present to identify a distribution of economic preference parameters (and associated theories) from a known prior distribution. The novel approach here uses Equivalence Class Determination as a way to to make inferences about noisy choice theories using methods developed for a classic, noiseless case. The trick is intuitively simple: A noiseless deterministic theory will make a single 0-1 choice in each choice pair. A small amount of noise is tolerated by "flipping" those 0-1 bits in a vector of choices and creating noise-injected "copies" of a target theory. Then the algorithm only tries to distinguish between an equivalence class (the target theory and all its noisy copies) and other equivalence classes.

We introduced an adaptive greedy (myopic) algorithm EC^2, to choose questions from a design space. The algorithm measures edge strength by the product of the probabilities of two theories with similar predictions, then chooses a test which will "cut" the strongest edge (by ruling out one theory in the edge-pair and letting the other survive). The edge-cutting objective function is adaptive submodular, which leads to nice theoretical properties.

We applied EC^2 to an experimental design problem choosing among six different theories of risky choice (expectational theories, such as EU and prospect theory, and moment-weighting theories). Ground-truth simulations, which evaluate how well the algorithm approximates behavior that is known (from simulation) shows that EC^2 chooses the right theories rapidly (within 5-20 trials), and more rapidly than other theories for modest error rates.

In our empirical application, the method indicates that most subjects choosing among monetary gambles are classified (with high posterior probability) as using prospect theory evaluation, or EV.

As economists, we are also naturally concerned that if subjects understand that the algorithm is adaptive, there is a strategic incentive to misreport preferences in early trials to create better choices in latter trials. However, we argue that computing the value of misrepresentation is difficult; The value is small even when it can be computed (e.g., 8% increase in earnings), and misrepresentation is likely to leave telltale signs in a time series of choices (high early response times and switching of posterior modes partway through the experiment) which we simply do not see in data. As a last check, in one experimental treatment we explained that the method of selecting choice questions was adaptive and that, in theory, the subjects could earn more if they exploited this property. Subjects in that treatment did not behave significantly differently and did not earn more money.

The EC^2 criteria is also applicable when more than 2 choices are presented, e.g. if subjects are asked to choose between one of 3 lotteries in every round. It can even be used in the case where subjects are asked to

provide an ordinal ranking over several choices. Both the cases increases the information obtained per round, potentially leading to increased accuracy with a fewer number of questions. The caveat is that is potentially complicates the decision for the subjects, or prevent their understanding of the instructions. In lab experiments with simple choices (Reutskaja *et al.* (2011)), too many choices has been shown to lead to biases in results due to choice overload: such as subjects searching more in a focal region of the display.

The major future direction for research is to implement these algorithms to efficiently sort through theories in many areas of empirical economics, where multiple theories have coexisted for a long time.

Appendix A: Details of the Experiment

User Interface

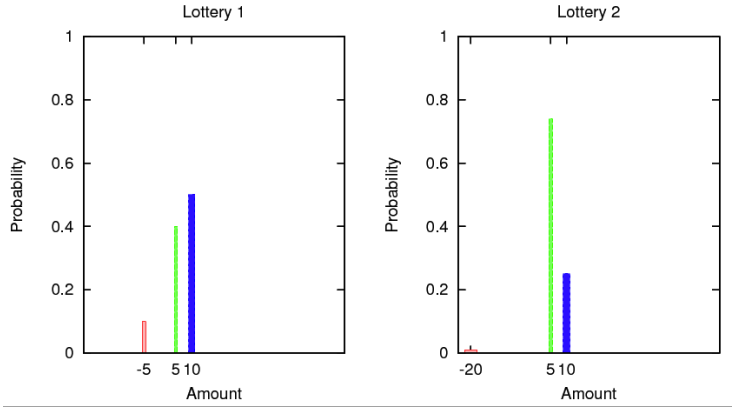


Table 2.5: *The User Interface presented to the participants. The two amounts and probabilities for the 2 lotteries are shown in text on the right. The same information is displayed as a graph on the left.*

Instructions

This is an experiment on the economics of decision making. If you follow the instructions and make good decisions you can earn a considerable amount of money which will be paid to you in cash. The funds have been supplied by a foundation that supports this research.

The experiment has 60 rounds. At the beginning you will begin with an endowment of $20. In each round you will choose between one of two lotteries.

Each lottery has 3 possible outcomes:

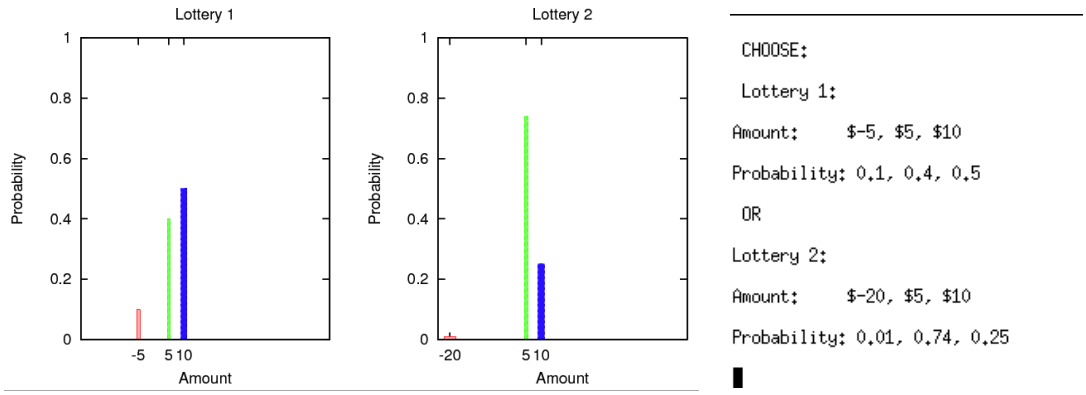
(i) One negative (loss) outcome of -$5, -$10 or -$20. (If you lose this amount it will be subtracted from your initial $20 endowment and you will be paid the remainder.)

(ii) One positive (gain) outcome of $5.

(iii) One positive (gain) outcome of $10, $20 or $50.

Each of the three possible outcomes has an associated probability of occurring. You will see a textbox with the two lottery choices described in text, on the right side of the screen. On the left side of the screen, you will see the same information displayed as a histogram. The height of each line represents the probability of that outcome occurring. The $5 outcome is always shown in green, the higher positive outcome in blue, and the negative outcome in red.

For example, you might see the following screen:



In Lottery 1, you can lose $5 with probability 0.1, gain $5 with probability 0.4 or gain $10 with probability 0.5.

In Lottery 2, you can lose $20 with probability 0.01, gain $5 with probability 0.74 or gain $10 with probability 0.25.

You must close the graph display by clicking on the X (close) button. Then you can enter your choice into the textbox by pressing 1 or 2.

There is no time limit for viewing the graph and entering your choice. After you enter your choice, the screen will say REST. Another choice between two lottery options will be displayed after a couple of seconds.

After finishing all 60 rounds, we will randomly pick one of the 60 pairs of lotteries you were asked about to determine your additional earnings. Each lottery pair is equally likely to be chosen so you should treat each one as if it could be the paid round.

To determine which of the three outcomes you receive a random integer number between 1 and 100 will determine the outcome you get. Each of those integers will be equally likely (using a procedure described next).

To use the example above, suppose you chose Lottery 1. In that lottery, the outcome -$5 has 10% probability. Therefore, if the random number chosen is 1-10 then the outcome will be -$5. Notice that the range 1-10 is 10% of all the numbers from 1 to 100. So the percentage of possible numbers that would give the outcome -$5 corresponds to the probability .10.

In the lottery 1 example, the outcome will be +$5 with probability .4, and +$10 with probability .5. To implement these probabilities fairly, numbers from 11 to 50 will lead to the outcome +$5 and numbers from 51-100 will lead to the outcome +$10. Those two ranges of numbers are 40% and 50% of all the numbers 1-100, so the percentage of possible numbers corresponds to the probabilities stated in the text and shown on

the graph.

Additional Instructions for Manipulability Check

These additional instructions appended to the original instructions for Manipulability Check condition:

It could help you to know that an adaptive algorithm generates the choices between two lotteries that are presented to you, in the first 30 lottery choices. That is, your previous choices determine the following questions that are presented to you, because the algorithm is trying to determine the kinds of risks you like to take and the kinds of risks you do not like to take. It is possible, mathematically, to choose lotteries you do not like in the beginning in order to lead the adaptive algorithm to create future lottery choices that you like better. If you think you can guide the algorithm in this way, by choosing lotteries you dont like to create bettery choice pairs in future trials, you are free to do so. However, your final payoff will be calculated from one of the questions selected at random.

Quiz before the Experiment

1. There is always a possible $5 outcome for every lottery (circle one)

   TRUE FALSE

2. The red lines in the graph show the larger amounts that could be won ($10, $20, $50) (circle one)

   TRUE FALSE

3. If the row I choose has a random number which leads to a loss -$20 for the lottery that I chose, I still get to keep the $20 endowment. (circle one)

   TRUE FALSE

Self-Report after Experiment

1. Did you think that the questions in the first part were generated by an adaptive algorithm (circle one)

   YES NO

2. Have you figured out how the adaptive algorithm chose the subsequent questions based on your responses? If so, can you state the strategy you used to game it?

3. What strategy did you use to maximize your potential earnings in the first round? Please write a few sentences.

4. How was your strategy in the first part different from your strategy in the second part? Please write a few sentences.

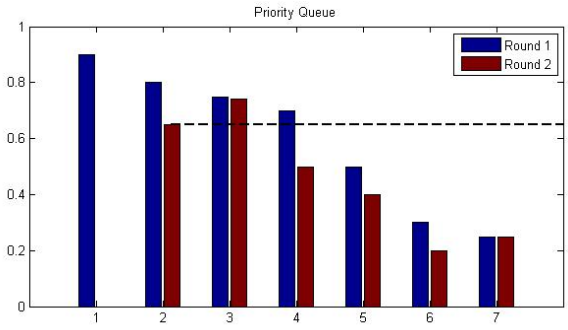


Figure 2.17: Sorted values of tests for the accelerated EC^2 implementation.

Appendix B

Details of the Efficient Implementation

Figure 2.17 illustrates the efficient accelerated implementation of EC^2 using a Priority Queue. Initially, the the tests are sorted according to their value, and the first test (number 1) is selected in the first round (and removed from the queue). In the second iterations, the value of the most valuable test (number 2) is recalculated, and only the tests that had a higher value than this in the previous iteration needs to be recalculated (number 3 and 4). The other tests (from 5) cannot have a higher value in round 2 than test number 2, and can be ignored. Algorithm 2 outlines the general procedure.

The methods were coded in Java SE 6.0. On a 2.53 GHz Intel computer with 4GB memory, the Information Gain criteria has to evaluate all 31,626 tests and takes 20.1 seconds to generate each question. The accelerated EC^2 method evaluates between 1 and up to 12 thousand tests at each iteration. By pre-computing all the values and efficient caching, the accelerated EC^2 method takes between 0.02 seconds to 6.1 seconds to generate a question. The average time is 0.4 seconds (over 1000 simulation runs of 100 rounds each).

Additional Proofs

Lemma 5. *The objective function f of Eq. (3.2.5) is strongly adaptive monotone.*

Proof. We must show that for all $\mathbf{x}_{\mathcal{A}}$, $t \notin \mathcal{A}$ and possible answer x for test t that

$$\mathbb{E}_H[f(\mathcal{A}, H) \mid \mathbf{x}_{\mathcal{A}}] \leq \mathbb{E}_H[f(\mathcal{A} \cup \{t\}, H) \mid \mathbf{x}_{\mathcal{A}}, X_t = x] \tag{2.8.1}$$

Towards this end, it is useful to notice that for all $t \in \mathcal{T}$ the function $h \mapsto \mathcal{E}_t(h)$ depends only on X_t. Hence for any $\mathbf{x}_{\mathcal{A}}$, the function $h \mapsto f(\mathcal{A}, h)$ is constant over realizations $\mathbf{x}_{\mathcal{T}} \succ \mathbf{x}_{\mathcal{A}}$, so we can define a function $g(\mathbf{x}_{\mathcal{A}})$ such that $g(\mathbf{x}_{\mathcal{A}}) = \mathbb{E}_H[f(\mathcal{A}, H) \mid \mathbf{x}_{\mathcal{A}}]$ by $g(\mathbf{x}_{\mathcal{A}}) := w\left(\bigcup_{t \in \mathcal{A}} \mathcal{E}_t(x_t)\right)$ where $\mathbf{x}_{\mathcal{A}} = (x_t)_{t \in \mathcal{A}}$ and

Input: Set of hypotheses $\mathcal{H}$; Set of tests $\mathcal{T}$; prior distribution P; function f.
begin
  $\mathcal{A} \leftarrow \emptyset$; $\Psi \leftarrow \emptyset$; PriorityQueue $Q \leftarrow$ EmptyQueue
  foreach $t \in \mathcal{T}$ **do**
    Q.insert($t, +\infty$)
  while $\exists h \neq h' : P(h \mid \mathbf{x}_\mathcal{A}) > 0$ *and* $P(h' \mid \mathbf{x}_\mathcal{A}) > 0$ **do**
    $\delta_{max} \leftarrow -\infty$; $t_{max} \leftarrow NULL$;
    while $\delta_{max} < Q.maxPriority()$ **do**
      $t \leftarrow Q.pop()$;
      $\delta \leftarrow \Delta(t|\Psi)$;
      $Q.insert(t, \delta)$;
      if $\delta_{max} < \delta$ **then**
        $\delta_{max} \leftarrow \delta$; $t_{max} \leftarrow t$;
    Set $\mathcal{A} \leftarrow \mathcal{A} \cup \{t_{max}\}$; $Q.remove(t_{max})$;
    Observe outcome $x_{t_{max}}$; Set $\Psi \leftarrow \Psi \cup \{(t_{max}, x_{t_{max}})\}$

Algorithm 2: Algorithm for accelerated EC^2. $Q.insert(t, \delta)$ inserts t with priority δ, $Q.pop()$ removes and returns the item with greatest priority, $Q.maxPriority()$ returns the maximum priority of the elements in Q, and $Q.remove(t)$ deletes t from Q.

$\mathcal{E}_t(x)$ is the set of edges cut by t if $X_t = x$. Note that for all $\mathbf{x}_\mathcal{A} \prec \mathbf{x}_\mathcal{B}$ we have $g(\mathbf{x}_\mathcal{A}) \leq g(\mathbf{x}_\mathcal{B})$, since the edge weights are nonnegative. Setting $\mathcal{B} = \mathcal{A} \cup \{t\}$ yields Eq. (2.8.1) and hence implies strong adaptive monotonicity. □

Lemma 6. *The objective function f of Eq. (3.2.5) is adaptive submodular for any prior with rational values.*

Proof. We first prove the result assuming a uniform prior $P(\cdot)$, and then show how to reduce the general prior case to the uniform prior case. Hence all edges have weight $1/n^2$, where there are n hypotheses. For convenience, we also rescale our units of reward so that all edges have unit weight. (Note that f is adaptive submodular iff cf is for any constant $c > 0$.) To prove adaptive submodularity, we must show that for all $\mathbf{x}_\mathcal{A} \prec \mathbf{x}_\mathcal{B}$ and $t \in \mathcal{T}$, we have $\Delta(t \,|\, \mathbf{x}_\mathcal{B}) \leq \Delta(t \,|\, \mathbf{x}_\mathcal{A})$. Fix t and $\mathbf{x}_\mathcal{A}$, and let $\mathcal{V}(\mathbf{x}_\mathcal{A}) := \{h : P(h \mid \mathbf{x}_\mathcal{A}) > 0\}$ denote the version space, if $\mathbf{x}_\mathcal{A}$ encodes the observed outcomes. Let $n_\mathcal{V} := |\mathcal{V}(\mathbf{x}_\mathcal{A})|$ be the number of hypotheses in the version space. Likewise, let $n_{i,a}(\mathbf{x}_\mathcal{A}) := |\{h : h \in \mathcal{V}(\mathbf{x}_\mathcal{A}, X_t = a) \cap \mathcal{H}_i\}|$, and let $n_a(\mathbf{x}_\mathcal{A}) := \sum_{i=1}^m n_{i,a}(\mathbf{x}_\mathcal{A})$. Also, define $e_a(\mathbf{x}_\mathcal{A}) := \frac{1}{2}\sum_{i \neq j}\sum_{b \neq a} n_{i,b}(\mathbf{x}_\mathcal{A}) \cdot n_{j,b}(\mathbf{x}_\mathcal{A})$ to be the number of edges cut such that at t both hypotheses agree with each other but disagree with the realized hypothesis h^*, conditioning on $X_t = a$. We define a function θ of these quantities such that $\Delta(t \,|\, \mathbf{x}_\mathcal{A}) = \theta(\mathbf{n}(\mathbf{x}_\mathcal{A}), \mathbf{e}(\mathbf{x}_\mathcal{A}))$, where $\mathbf{n}(\mathbf{x}_\mathcal{A})$ is the vector consisting of $n_{i,a}(\mathbf{x}_\mathcal{A})$ for all i and a and $\mathbf{e}(\mathbf{x}_\mathcal{A})$ is the vector consisting of $e_a(\mathbf{x}_\mathcal{A})$ for all a. For brevity, we suppress the dependence of $\mathbf{x}_\mathcal{A}$ where it is unambiguous. Then, as we will explain below, θ is defined as

$$\theta(\mathbf{n}, \mathbf{e}) := \frac{1}{2}\sum_{i \neq j}\sum_{a \neq b} n_{i,a} \cdot n_{j,b} + \sum_a e_a \left(1 - \frac{n_a}{n_\mathcal{V}}\right) \tag{2.8.2}$$

Here, i and j range over all class indices, and a and b range over all possible outcomes of test t. The first

term on the right-hand side counts the number of edges that will be cut by selecting test t no matter what the outcome of t is. Such edges consist of hypotheses that disagree with each other at t and, as with all edges, lie in different classes. The second term counts the expected number of edges cut by t consisting of hypotheses that agree with each other at t. Such edges will be cut by t iff they disagree with the true hypothesis h^* at t. The edges $\{h, h'\}$ with $h, h' \in \mathcal{V}(\mathbf{x}_\mathcal{A})$ and $P(X_t = a \mid h) = P(X_t = a \mid h') = 1$ (of which there are e_a) will be cut by t iff $X_t \neq a$. Since we assume a uniform prior, $\mathbb{P}\left[X_t \neq a \;\middle|\; \mathbf{x}_\mathcal{A}\right] = 1 - n_a/n_\mathcal{V}$ for any partial realization $\mathbf{x}_\mathcal{A}$ with $t \notin \mathcal{A}$, hence the expected contribution of these edges to $\Delta\left(t \,|\, \mathbf{x}_\mathcal{A}\right)$ is $e_a\left(1 - n_a/n_\mathcal{V}\right)$, from whence we get the second term.

Now fix $\mathbf{x}_\mathcal{B} \succ \mathbf{x}_\mathcal{A}$. Our strategy for proving $\Delta\left(t \,|\, \mathbf{x}_\mathcal{B}\right) \leq \Delta\left(t \,|\, \mathbf{x}_\mathcal{A}\right)$ is as follows. As more observations are made, the version space can only shrink, i.e. $\mathcal{V}(\mathbf{x}_\mathcal{B}) \subseteq \mathcal{V}(\mathbf{x}_\mathcal{A})$. This means that for all i and a, $n_{i,a}$ is nonincreasing, i.e., $n_{i,a}(\mathbf{x}_\mathcal{B}) \leq n_{i,a}(\mathbf{x}_\mathcal{A})$. Note we may interpret e_a as a function of the variables in $\{n_{i,a} : 1 \leq i \leq m, a \in \mathcal{X}\}$, and that it is nondecreasing in each $n_{i,a}$, so we may also deduce that $e_a(\mathbf{x}_\mathcal{B}) \leq e_a(\mathbf{x}_\mathcal{A})$ for all a. Hence we consider a parameterized path $p(\tau)$ in $\mathbb{R}^{(m+1)\cdot\ell}$ from $p(0) := (\mathbf{n}(\mathbf{x}_\mathcal{B}), \mathbf{e}(\mathbf{x}_\mathcal{B}))$ to $p(1) := (\mathbf{n}(\mathbf{x}_\mathcal{A}), \mathbf{e}(\mathbf{x}_\mathcal{A}))$. Then by integrating along the path we obtain

$$\Delta\left(t \,|\, \mathbf{x}_\mathcal{A}\right) - \Delta\left(t \,|\, \mathbf{x}_\mathcal{B}\right) = \int_{\tau=0}^{1} \left(\frac{d(\theta \circ p)}{d\tau}\right) d\tau. \tag{2.8.3}$$

We require that at each point in p it holds that $e_a = \frac{1}{2}\sum_{i \neq j} \sum_{b \neq a} n_{i,b} \cdot n_{j,b}$ for all a, and also ensure that p is nondecreasing in each coordinate. There exists a path meeting these requirements, since $(\mathbf{n}(\mathbf{x}_\mathcal{B}), \mathbf{e}(\mathbf{x}_\mathcal{B})) \leq (\mathbf{n}(\mathbf{x}_\mathcal{A}), \mathbf{e}(\mathbf{x}_\mathcal{A}))$ and each e_a is nondecreasing in each $n_{i,b}$ variable. This implies $\partial n_{i,a}/\partial\tau \geq 0$ and $\partial e_a/\partial\tau \geq 0$ for all i and a. Hence we can prove the integral is nonnegative by applying the chain rule for the derivative to obtain

$$\frac{d(\theta \circ p)}{d\tau} \;=\; \sum_{i,a} \frac{\partial\theta}{\partial n_{i,a}} \frac{\partial n_{i,a}}{\partial\tau} + \sum_{a} \frac{\partial\theta}{\partial e_a} \frac{\partial e_a}{\partial\tau}$$

and then proving that $\partial\theta/\partial n_{i,a} \geq 0$ and $\partial\theta/\partial e_a \geq 0$ for all i and a. Next, observe that $\partial\theta/\partial e_a = (1 - n_a/n_\mathcal{V}) \geq 0$. So fix a class index k and an outcome c and consider $\partial\theta/\partial n_{k,c}$. Elementary calculus tells us that

$$\frac{\partial\theta}{\partial n_{k,c}} \;=\; \sum_{j \neq k,\; b \neq c} n_{j,b} \;+\; \sum_{b} \frac{e_b n_b}{n_\mathcal{V}^2} \;-\; \frac{e_c}{n_\mathcal{V}} \tag{2.8.4}$$

This quantity is nonnegative iff

$$n_\mathcal{V} e_c \;\leq\; n_\mathcal{V}^2 \cdot \sum_{j \neq k,\; b \neq c} n_{j,b} \;+\; \sum_{b} e_b n_b \tag{2.8.5}$$

Now substitute $\frac{1}{2}\sum_{i\neq j}\sum_{b\neq c} n_{i,b}\cdot n_{j,b}$ for e_c to obtain

$$n_{\mathcal{V}} e_c = \frac{n_{\mathcal{V}}}{2}\sum_{i\neq j}\sum_{b\neq c} n_{i,b}\cdot n_{j,b} \leq n_{\mathcal{V}}\cdot\left(\sum_{j\neq k,\ b\neq c} n_{j,b}\right)\left(\sum_{i,a} n_{i,a}\right) \leq n_{\mathcal{V}}^2\cdot\left(\sum_{j\neq k,\ b\neq c} n_{j,b}\right) \tag{2.8.6}$$

Since $\sum_b e_b n_b \geq 0$, we obtain Eq. (2.8.5) from Eq. (2.8.6) by inspection, and hence $\partial\theta/\partial n_{k,c} \geq 0$ for all k and c. This completes the proof of the adaptive submodularity of f under a uniform prior.

We now show how to reduce the general prior case to the uniform prior case. Fix any prior P with rational probabilities, i.e. $P(h) \in \mathbb{Q}$ for all h. Then there exists $d \in \mathbb{N}$ and function $k : \mathcal{H} \to \mathbb{N}$ such that such that $P(h) = k(h)/d$. Create a new instance containing d hypotheses, where for each $h \in \mathcal{H}$ there are $k(h)$ copies of h, denoted by $h^1, \ldots, h^{k(h)}$. Each copy of h induces the same conditional distribution of test outcomes $P(X_1, \ldots, X_N \mid h)$. All copies of h belong to the same class, and copies of h and h' belong to the same class iff h and h' do. Finally, assign a uniform prior to this new instance. Then the adaptive submodularity of f on this new instance implies the adaptive submodularity on the original instance, if the weight of edge $\{h, h'\}$ in the original instance is proportional to the number of edges between the copies of h and the copies of h' in the new instance. That is, it suffices to set $w(\{h, h'\}) \propto k(h)\cdot k(h')$, and our choice of weight function, $w(\{h, h'\}) := P(h)\cdot P(h')$, satisfies this condition. □

A Bad Example for Information Gain

A popular heuristic for the Optimal Decision Tree problem are to adaptively greedily select the test that maximizes the *information gain* in the distribution over hypotheses, conditioned on all previous test outcomes. The same heuristic can be applied to the Equivalence Class Determination problem, in which we compute the information gain with respect to the entropy of the distribution over *classes* rather than hypotheses. Let π_{IG} denote the resulting policy for Equivalence Class Determination.

Another common heuristic for Optimal Decision Tree is to adaptively greedily select the test maximizing the *Bayesian decision-theoretic value of information* (VoI) criterion. Recall the value of information of a test t is the expected reduction in the expected risk of the minimum risk decision, where the risk is the expected loss. Formally, consider the Bayesian decision-theoretic setup described in §2.3.3. The VoI criterion myopically selects test to maximize

$$\Delta_{\text{VoI}}(t \,|\, \mathbf{x}_{\mathcal{A}}) := \min_d \mathbb{E}_H[\ell(d, H) \mid \mathbf{x}_{\mathcal{A}}] - \mathbb{E}_{x_t \sim X_t|\mathbf{x}_{\mathcal{A}}}\left[\min_d \mathbb{E}_H[\ell(d, H) \mid \mathbf{x}_{\mathcal{A}}, x_t]\right].$$

This heuristic can be also be applied to the Equivalence Class Determination problem, by taking the decision

set $\mathcal{D}$ to be the set of equivalence classes, and the loss function to be the 0–1 classification loss function, i.e., $\ell(d, H) = \mathbf{1}[H \notin d]$. Let π_{VoI} denote the resulting policy.

In this section we present a family of Equivalence Class Determination instances for which both π_{IG} and π_{VoI} perform significantly worse than the optimal policy.

Theorem 7. *There exists a family of Equivalence Class Determination instances with uniform priors such that $c(\pi_{IG}) = \Omega\left(n/\log(n)\right) c(\pi^*)$ and $c(\pi_{VoI}) = \Omega\left(n/\log(n)\right) c(\pi^*)$, where n is the number of hypotheses and π^* is an optimal policy.*

In fact, we will prove a lower bound for each policy within a large family of adaptive greedy policies which contains π_{IG} and π_{VoI}, which we call *posterior–based*. Informally, this family consists of all greedily policies that use only information about the posterior equivalence class distribution to select the next test. More precisely, these policies define a potential function Φ which maps distributions of distributions over equivalence classes to real numbers, and at each time step select the test t which maximizes Φ of the posterior distribution (over test outcomes x_t) of the posterior distribution over equivalence classes generated by adding x_t to the previously seen test outcomes. In the event of a tie, we select any test maximizing this quantity at random. The information gain policy is posterior–based; Φ is simply -1 times the expected entropy of the posterior equivalence-class distribution. Likewise, the value of information policy is also posterior–based; Φ is simply -1 times the expected loss of the best action for the posterior equivalence-class distribution. Hence to prove Theorem 7 it suffices to prove the following more general theorem.

Theorem 8. *There exists a family of Equivalence Class Determination instances with uniform priors such that $c(\pi) = \Omega\left(n/\log(n)\right) c(\pi^*)$ for any posterior–based policy π, where n is the number of hypotheses and π^* is an optimal policy.*

Proof. Fix integer parameter $q \geq 1$. There are $m = 2^q$ classes $\mathcal{H}_a$ for each $1 \leq a \leq 2^q$. Each $\mathcal{H}_a$ consists of two hypotheses, $h_{a,0}$ and $h_{a,1}$. We call a the *index* of $\mathcal{H}_a$. The prior is uniform over the hypotheses $\mathcal{H} = \{h_{a,v} : 1 \leq a \leq m, 0 \leq v \leq 1\}$. There are four types of tests, all with binary outcomes and all of unit cost. There is only one test of the first type, t_0, which tells us the value of v in the realized hypothesis $h^*_{a,v}$. Hence for all a, $H = h_{a,v} \Rightarrow X_{t_0} = v$. Tests of the second type are designed to help us quickly discover the index of the realized class via binary search if we have already run t_0, but to offer no information gain whatsoever if t_0 has not yet been run. There is one such test t_k for all t with $1 \leq k \leq q$. For $z \in \mathbb{N}$, let $\phi_k(z)$ denote the k^{th} least-significant bit of the binary encoding of z, so that $z = \sum_{k=1}^{\infty} 2^{k-1}\phi_k(z)$. Then for each $h_{a,v}$ we have $H = h_{a,v} \Rightarrow X_{t_k} = \mathbf{1}[\phi_k(a) = v]$. Tests of the third type are designed to allow us to do a (comparatively slow) sequential search on the index of the realized class. Specifically, we have tests t_k^{seq} for all

$1 \leq k \leq m$, such that $H = h_{a,v} \Rightarrow X_{t_k^{\text{seq}}} = \mathbf{1}[a = k]$. Finally, tests of the fourth type, $\left\{t_k^{\text{dumb}} : k \in \mathbb{N}\right\}$, are dummy tests that reveal no information at all. Formally, $X_{t_k^{\text{dumb}}}$ always equals zero.

Given this input, suppose $H = h_{a,v}$. One solution is to run t_0 to find v, then run tests $t_1, \ldots, t_q$ to determine $\phi_k(a)$ for all $1 \leq k \leq q$ and hence to determine a. This reveals the value of H, and hence the class H belongs to. Since the tests have unit cost, this policy π' has cost $c(\pi') = q + 1$.

Next, fix a posterior–based policy π and consider what it will do. Call a class *possible* if not all of its hypotheses have been ruled out by tests performed so far. Note that all possible classes contain the same number of hypotheses, because they initially have two, and each test t_k that can reduce the size of a possible class to one, will reduce the size of every possible class to one. This, and the fact that the prior is uniform, implies that the posterior equivalence-class distribution is uniform over the remaining possible classes. If no tests in $\{t_k : 0 \leq k \leq q\}$ have been run, as is initially the case, any single test in this set will not change the posterior equivalence-class distribution. Hence, as measured with respect to Φ, such tests are precisely as good as the dummy tests. If these tests are each better than any test in $\left\{t_k^{\text{seq}} : 1 \leq k \leq m\right\}$, then π selects among $\{t_k : 0 \leq k \leq q\} \cup \left\{t_k^{\text{dumb}} : k \in \mathbb{N}\right\}$ at random. Since there are infinitely many dummy tests, with probability one a dummy test is selected. Since the posterior remains the same, π will repeatedly select a test at random from this set, resulting in an infinite loop as dummy tests are selected repeatedly *ad infinitum*. Otherwise, some test t_k^{seq} is preferable to the other tests, measured with respect to Φ. In the likely event that t is not the index of H, we are left with a residual problem in which tests in $\{t_k : 0 \leq k \leq q\}$ still have no effect on the posterior, there is one less class, and the prior is again uniform. Hence our previous argument still applies, and π will either enter an infinite loop or will repeatedly select tests in $\left\{t_k^{\text{seq}} : 1 \leq k \leq m\right\}$ until a test has an outcome of 1. Thus in expectation π costs at least $c(\pi) \geq \frac{1}{m}\sum_{z=1}^{m} z = (m+1)/2$. Since $m = 2^q$, $n = 2m$, and $c(\pi^*) \leq c(\pi') = q + 1 = \log_2(n)$ we infer

$$c(\pi) \geq \frac{m}{2} = \left(\frac{n}{4\log_2(n)}\right) c(\pi^*)$$

which completes the proof. □

Chapter 3

Time Preference

The rational prescription for making decisions that are delayed in time uses an exponential discounting function, and avoids temporal inconsistency (Strotz (1955)). Now, there is a growing literature in experimental and behavioural economics providing evidence that human behaviour is indeed time-inconsistent, and people are willing to forego larger delayed rewards for smaller immediate rewards (see Frederick *et al.* (2002); Doyle (2013), for literature surveys). Economic models that account for this departure from rationality vary from the one-parameter hyperbolic discounting function, to *present-bias* models (e.g. quasi-hyperbolic discounting and fixed time cost models that have an additional parameter) that account for the observation that people pay a premium to choose options that are immediately available. Two-parameter, generalized-hyperbolic discounting models (Loewenstein and Prelec (1992)) nest exponential discounting models, and hyperbolic discounting models, for specific parameter values.

Time preference (also called temporal discounting) models have been studied in many contexts, in both economics and psychology. In consumer behaviour, Estle *et al.* (2007) looked at monetary and consumable rewards (and found steeper discounting for the latter); and rewards as leisure time (Chapman and Elstein (1995)). Others have looked at drug addiction, smoking, gambling and alcoholism and their relationship to delay discounting (Madden and Bickel (2010)).

Despite the multitude of models, and experiments with them in different contexts, the literature containts few face-to-face comparisons. Furthermore, since time preference seems to be highly dependent on the domain or context (Hardisty *et al.* (2012); Tsukayama and Duckworth (2010); Weatherly *et al.* (2010)), a universal discounting model may not be appropriate. A framework for efficiently comparing time preference models can inform decision-makers in choosing the appropriate model to use within a context.

Choices used in experimental and survey-based social science are typically developed by hunches and cumulative searches for informative questions to test theories. The conventional designs that have emerged are

typically a fixed set of test questions .[1] We apply an approach, developed previously in Ray *et al.* (2013), in which the sequence of choices is tailored for each subject rather than fixed. The subjects themselves tell us, through their answers, the "best"(most informative) question to ask them next.

Our approach, **B**ayesian **R**apid **O**ptimal **A**daptive **D**esign) (abbreviated as BROAD), is an innovation in a large family of adaptive methods (described in section 3.2). The key innovation is an edge-cutting measure of information value that is adaptively submodular, which therefore provably guarantees some useful theoretical and practical properties. The method was introduced by Golovin *et al.* (2010), Ray *et al.* (2013), and applied here to novel economic questions.

Earlier applications of adaptive methods were made in statistics (Lindley (1956)), decision theory (Howard (1966)), computer-assisted testing (CAT) in psychometrics (Wainer (1990)), cognitive psychology (Myung and Pitt (2009)), adaptive choice-based conjoint measurement in marketing (Abernethy *et al.* (2007)), active learning methods in machine-learning (Golovin and Krause (2010b); Nowak (2009); Dasgupta (2004)), and management science (Cavagnaro *et al.* (ming); Toubia *et al.* (ming)).

Some early efforts to introduce optimal adaptive design in experimental economics (e.g., El-Gamal *et al.* (1993), El-Gamal and Palfrey (1993), Moffatt (2007)) did not gain traction. The time is now ripe for BROAD methods because: Computing power is better than ever; a new method from computer science (called EC^2) applied here provides theoretical guarantees on efficient computability; and there are many new competing theories in behavioral and experimental economics which need to be efficiently compared.

In our application, and all the others mentioned above, the goal is to sequentially select among a set of noisy, expensive observations (outcomes of experiments, medical tests to run to diagnose patients, etc.) in order to determine which hypothesis (theory, diagnosis, etc.) is most accurate. One way to formalize such active learning problems is *Bayesian experimental design* (Chaloner and Verdinelli (1995)). This approach specifies prior beliefs over a set of hypotheses, as well as probabilistic assumptions about the outcomes of test choices. The goal then is to determine the correct hypothesis while minimizing the cost of the experimentation (where cost is synonymous, for present purposes, with the number of test choice questions that are asked). Unfortunately, finding an optimal sequence of test choices is not just computationally difficult (NP-hard) but is also difficult to approximate (Chakaravarthy *et al.* (2007)). Several heuristic approaches have been proposed that perform well in some specific applications, but do not have theoretical guarantees (e.g., MacKay (1992)); that is, there are no proofs about how costly the heuristic sequence will be compared to the optimal sequence.

In the case where observations are *noise-free*[2], a simple algorithm, *generalized binary search*[3]*(GBS)*, is

[1] Note that "test questions", in experimental economics could include choices from budget sets Andreoni and Miller (2002); Choi *et al.* (2007); Fisman *et al.* (2007), strategies in games, auction bids, and trading strategies. Our method could also be applied to large-scale panel surveys in which branching methods are already used.

[2] This case is known as the *Optimal Decision Tree* (ODT) problem.

[3] GBS selects test choices to maximize, in expectation over the test outcomes, the probability mass of eliminated hypotheses (i.e., those

guaranteed to be competitive with the optimal test sequence[4]. Unfortunately, when responses are likely to be imperfect or *noisy*, as in economics experiments, the theoretical basis and performance of heuristic methods is not well understood.[5]

In this paper, we introduce and apply a general method, called BROAD, using a formulation of Bayesian active learning with noisy observations that we call the *Equivalence Class Determination* problem. BROAD utilizes a myopic (greedy) active learning algorithm using the **E**quivalence **C**lass **E**dge **C**utting (EC^2) objective function, that can be proved to have an expected cost that is a multiple of the optimal test sequence cost. The key insight is that the EC^2 objective function for measuring information value satisfies *adaptive submodularity* (Golovin and Krause (2011)). Adaptive submodularity is a natural diminishing returns property that generalizes the classical notion of submodularity to adaptive policies, and leads to a provable performance guarantee as well as a shortcut in evaluating test choices (called "lazy evaluation").

Here are some possible *advantages* of BROAD approaches:

○ The posterior distribution of all theory and parameter probabilities is quickly recomputed for each subject after each trial (because it is a necessary step in finding the optimal next question). As a result much of the data analysis is already done when the experiment is over.

○ Because of that instant probability updating, BROAD creates a statistical parametric group portrait after each trial. The portraits show who seems most impatient, most averse to risk, most reciprocal, etc. These data could then be used to instantly form groups of different types of people to see how those differences affect group or market behavior.

○ Since BROAD techniques economize on information gained per minute, they are especially useful for subject pools who have scarce time or become bored or habituated quickly. Such groups include highly-trained professionals, internet subjects who can quit, human groups such as lesion patients or children, and animals that typically make long sequences of lab choices.

○ The fact that the BROAD procedure generates sequences of test questions that are provably near-optimal can sharpen discourse about what different experimental designs are good and bad for. Novel BROAD designs which are unconventional should gain credibility if they have desirable informational properties. BROAD

with zero posterior probability, computed with respect to the observed test outcomes). The method is myopic (or "greedy" in computer science terms) because it only searches ahead for one test at a time.

[4]The expected number of tests is a factor of $O(\log n)$ (where n is the number of hypotheses) more than that of the optimal policy (Kosaraju *et al.* (1999)), which matches lower bounds up to constant factors (Chakaravarthy *et al.* (2007)). In other words, unless P = NP, no efficient algorithm can be guaranteed to find a policy whose expected number of tests is less than C log(n) times that of the optimal policy, for some absolute constant C (see Papadimitriou (2003)).

[5]There are no efficient algorithms that provably achieve minimal cost, except in very restricted settings (Krause and Guestrin (2009)). While there are some recent positive results in understanding the label complexity (bounds on the sample size required) of noisy active learning (Nowak (2009); Balcan *et al.* (2006)), the results depend on the assumption that the same query can be made multiple times with independent noise. This is not appropriate in our setting; if a test subject makes a mistake in evaluating a pair of gambles, presenting the exact same choice later on is likely to elicit the same mistake, which induces correlation in the noise across tests.

methods can also be used to judge the quality of older conventional designs[6].

○ At each trial it is easy to compute how much extra parameter precision (or model selection accuracy) is expected from asking one more question. This feature permits a cost-benefit method for optimally stopping the experiment (Swartz and Choi (2009)), if information benefit and cost can be compared.

○ In theory, subjects might prefer to strategically manipulate their early responses in order to get "better" (more economically valuable) future test questions. Ray *et al.* (2013) discussed tests to detect and minimize it in applications of BROAD, and found no evidence for strategic manipulation in experimental data.

In section 3.1, we give a brief overview of the time preference models that we compare in this paper. In section 3.2, we introduce the Bayesian Rapid Optimal Adaptive Design (BROAD) method, followed by implementation details and experimental setup in 3.3. We analyze the results of laboratory experiments comparing the time preference models in section 3.4, followed by a discussion of future applications of BROAD in time preference in section 3.5.

[6]For example, retrospective analysis of influential psychology experiments on memory retention indicates that some early designs were remarkably near-optimal but others were not (Myung and Pitt (2009))

3.1 Models

Formally, pairs (x, t) represent dollars x obtained at time t (equivalently, t periods from now). Assuming the subject's preferences over monetary rewards are linear, the discount function $D(x, t)$ is defined so that the subject is indifferent between the pair (x, t) and the pair $(xD(x, t), 0)$. We say that "the value of x at time t is $xD(x, t)$.

The standard method in finance of translating present values into future values is by a geometric mean taken on the ratio of increase[7]. As the compounding becomes continuous, we get the *Exponential* discounting function with the form:

$$D(x, t) = \exp(-rt), \quad r > 0 \tag{3.1.1}$$

The exponential discounting function is also a normative discounting model that produces time-consistent choices.

In accounting, the arithmetic average rate of return is frequently calculated (Ryan (2007)). While a model of compound interest underpins exponential discounting, hyperbolic discounting is underpinned by a model of simple interest, with no compounding (Rachlin (2006)). *Hyperbolic* discounting was also featured early on as a model of animal foraging (Mazur (1987)), with a functional form that induces declining subjective discount rates.

$$D(x, t) = \frac{1}{1 + rt}, \quad r > 0 \tag{3.1.2}$$

Quasi-hyperbolic discounting displays a *present bias*, and was introduced by Laibson (1997). The parameters $\theta = \{r, \alpha\}$ includes the present bias term α which induces a variable cost associated to future payoffs. The discount function $D(x, t)$ is given by:

$$D(x, t) = \begin{cases} 1, & \text{if t} = 0 \\ \alpha \exp(-rt), & \text{if t} > 0 \end{cases} \tag{3.1.3}$$

For $\alpha = 1$ we would have a straightforward compounding model, but if $\alpha < 1$, this parameters gives a one-off boost to discounting over the first compounding period. Quasi-hyperbolic discounting is only hyperbolic in mimicing the initial steep discounting of the hyperbolic model, but is an exponential model in every other sense.

Instead of a percentage decrement for non-present rewards, Benhabib *et al.* (2010), suggested using a fixed

[7]The compounding growth model for discrete time is, $F = P(1 + r)^T$, where F is the future value, P is the present value, r the growth rate, and T is the compounding periods. If we increase the number of compounding periods to n per annum, we get $F = P(1 + r/n)^{nT}$. For continous compounding, we get the exponential discounting form.

cost to model the present bias, that is, an absolute decrement. Any reward x is valued at $x - b$ when received in the future. The discount curve is:

$$D(x,t) = \begin{cases} 1, & \text{if t} = 0 \\ \exp(-rt) - \dfrac{b}{x}, & \text{if t} > 0 \end{cases} \tag{3.1.4}$$

Additionally, we consider the class of *Generalized hyperbolic* functions, first introduced in Loewenstein and Prelec (1992) and Prelec (2004) as the functional form:

$$D(x,t) = (1 + \alpha t)^{-(\theta/\alpha)}, \quad \alpha, \theta > 0. \tag{3.1.5}$$

The α coefficient determines how much the function departs from constant discounting. The limiting case, as α goes to zero, is the exponential discount function. When $\theta/\alpha = 1$, we get hyperbolic discounting; And when $\theta/\alpha < 1$, the function discounts more steeply than exponential at short durations, but less so at longer durations. McKerchar *et al.* (2010, 2009), found that modeling the additional parameter leads to statistically significant improvements in fit over the hyperbolic discounting model.

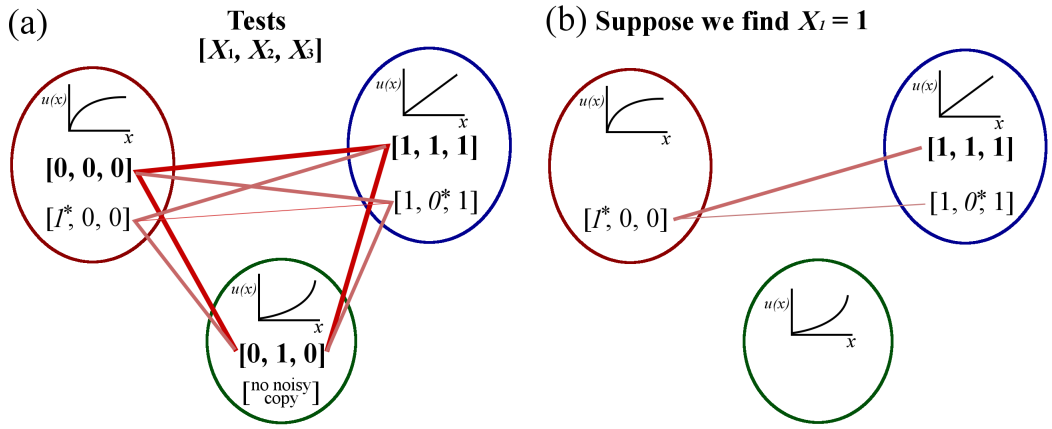


Figure 3.1: A stylized example of Equivalence Class Edge Cutting.

3.2 Optimal Adaptive Design

We begin with a brief history of adaptive methods, followed by an intuitive sketch of the objective function behind BROAD, and a theoretical outline. The appendix covers the theory in greater detail; interested readers should see Ray *et al.* (2013) for detailed proofs.

3.2.1 A brief history of related adaptive methods

The idea of optimizing an experimental design to measure parameters or test theories originated with Charles S. Peirce (Peirce (1967)), but did not have immediate influence. Dynamic design ideas began with Wald's influential sequential probability ratio test (Wald (1949)). Later contributions were made by Kiefer (1959), Atkinson and Fedorov (1975), and reviewed by Atkinson and Donev (1992) and Chaloner and Verdinelli (1995). D-optimal designs are optimal for estimating parameter values for a single theory. T-optimal designs optimally discriminate different theories and are more computationally challenging (cf. "landscaping", by D.J. Navarro and Myung (2004)). Many of the principles originally developed under D- and T-optimality now permeate standard experimental design practice.

The idea of adaptively customizing test choices goes back 100 years to Binet and Simon (1904). It was pushed forward decades later by the US Army (Bayroff *et al.* (1960)), and in psychometrics (Lord (1974); Weiss (1975); Weiss and Betz (1973)) in the 1960s and 70s.

Adaptive methods have also been used in neurophysiology (Lewi *et al.* (2009)), psychophysics (Kujala and Lukka (2006); L.A. Lesmes and Dosher (2006)), and medicine (Muller *et al.* (2007)). In marketing, adaptive choice-based "conjoint measurement" (ACBC) is used to discover which feature configurations of products consumers like most e.g., Sawtooth software (Johnson *et al.* (2003)), and in preference elicitation (Abernethy

et al. (2007); Toubia *et al.* (2004)). Improved methods will undoubtedly improve the ability to predict real behavior from responses which are hypothetical (often necessarily so). For example, Ding *et al.* (2005), Ding (2007), and Dong *et al.* (2010) describe incentive-aligned conjoint procedures. These methods use hypothetical data from conjoint measurements or product ranks combined with random selection of a single real product for purchase (based on WTP or rank inferred from conjoint or ranks), to boost incentive-compatibility of the conjoint responses. In cognitive psychology, adaptive methods have been championed by Myung and Pitt (2009), Myung *et al.* (2009), Cavagnaro *et al.* (2011), Cavagnaro *et al.* (2010, ming).

Two applications closely related to ours are Wang *et al.* (2010b); Toubia *et al.* (ming). Wang *et al.* (2010b) chooses questions that maximize Information-Gain, and Toubia *et al.* (ming) picks questions that maximize the norm of the Hessian of the posterior distribution at a point estimate (MLE) of the mode. None of the method have theoretical performance guarantees, while BROAD (Ray *et al.* (2013)) provides the first theoretical bounds when responses are noisy. Furthermore, we use an accelerated greedy implementation of EC^2 that generates new questions in real-time, without the need to store the question sequence in advance, thus there are no practical limitations on the number of rounds of questions asked.

We begin with intuition about EC^2, then describe the method. The Appendix has a more detailed development; but readers should see Ray *et al.* (2013) for proofs. It has three powerful properties which other information criteria do not have:

∘ Most methods are sensitive to noise.[8] The empirical applications which we have an mind all feature responses with some degree of stochastic noise (i.e., people choosing according to decision rule will not make the same choice every single time from a fixed choice set). Many of the good properties of previous adaptive methods rely on deterministic, no-noise processes (so that hypotheses are immediately eliminated by a single unpredicted "mistake"). EC^2 reduces the noisy case to a formally tractable noiseless case, so good performance properties carry over while noise is allowed.

∘ There is a provable worst-case cost bound for greedy EC^2 which depends on $ln(1/p_{min})$ (where p_{min} is the prior probability of the least-likely theory). What does that imply? In a medical setting, for example, there could be a dangerous disease which is extremely rare (p_{min} is low) but highly contagious; and figuring out for sure whether a patient has the disease is necessary for protecting public health. This kind of diagnosis will take many tests in the worst case. However, in the empirical social science settings we are considering, the lowest p_{min} is rarely very low because it is usually adequate to consider only a modest number of possible theories and parameters (p_{min} might be on the order of 10^{-3}, for example, so that the worst-case cost multiplier is 6.9).

∘ Since our EC^2 information criteria is adaptively submodular, the marginal information value of a test

[8] Ray *et al.* (2013) illustrates an example where information gain and value of information (VoI) perform badly.

can only go down after other tests have been done. This permits "lazy evaluation" which reduces runtime (Leskovec *et al.* (2007); Minoux (1978)).[9]

Next we will provide a simple visual illustration to develop intuition, followed by a more formal theoretical development (refer to Appendix for the details).

3.2.2 A simple illustration

Intuitively, EC^2 chooses the test which (a) distinguishes two hypotheses, connected by an edge, which make different predictions on that test, and (b) has the largest product of those two hypotheses' probability. The method is like a sports tournament in which all pairs of challengers that make different predictions are created (edges). Pairs are ranked by the product of their probabilities of being the "best" (i.e., most likely). The pair of contending theories with the highest probability product are then chosen to compete on a conclusive test (which will knock out one and be consistent with the other; In fact, a challenge question can maximize the sum of probability products for multiple edges, hence knocking out multiple theories at the same time). The idea is to choose the challenge which knocks out one of two theories –in our context, one of two theories that have the highest product of likelihoods of being true– as quickly as possible.

Figure 3.1a shows a graphic illustration of how EC^2 works. Each oval represents an equivalence class (EC) of theory predictions. In the upper left, the theory predicts choices of [0,0,0] for three tests X_1, X_2, X_3. A "noisy copy" of the theory's prediction [1*,0,0] is randomly generated by flipping the first choice from 0 to 1. (No noisy copy is created for the bottom oval theory.) If the noise rate is 10%, then the probability of observing choices [1*,0,0] is .10 times the probability of the true predicted choices [0,0,0]. Edges connect all pairs of predictions between equivalence sets (although, importantly, *no edges* connect *within* an equivalence set since those predictions all stem from noisy copies of the same true theory). The thickness of the edges is a visual index of the product of the probabilities of the predictions connected by that edge. The test which can cut the sum of the thickness of multiple edges is chosen.

Suppose the test X_1 is chosen. It can distinguish the true (noiseless) predictions in left red EC and in lower green EC, which predict 0 in X_1, from the upper right blue EC which predicts 1. If the choice is actually 1, then all theories that predicted 0 are eliminated and their edges disappear, as shown in Figure 3.1b.

An important caveat: Note that when implemented, the EC^2 algorithm applied to the ECD will not always choose the true hypothesis. Here's an example illustrating why: Suppose there are 10 binary tests. A

[9]To illustrate, suppose there are 30,000 possible tests. The procedure begins by computing the value of all 30,000 tests. The highest-value test is conducted on trial 1, the result is recorded, and all model probabilities are updated. On trial 2, if a candidate new test has a computed value V, then all the tests with a value lower than V in the initial pre-trial 1 evaluation can be safely ignored (since those values may change after the trial 1 results, but can never become larger than V). This insight can be exploited, by considering tests in order of maximum previously computed values (Refer to Ray *et al.* (2013) for implementation details).

hypothesis h_1 is a vector of ten 0's and 1's. Now suppose we flip those numbers so that all possible vectors are generated (i.e., there are 10 one-flip noisy copies, 10!/(10-2)!2=45 two-flip noisy copies, etc.). If we do that for another hypothesis h_2, then one of the noisy h_1 copies will exactly match one of the h_2 copies. We cannot draw an edge between them, because no test can distinguish them. The approach can't guarantee that we will pick the true hypothesis even if we conduct all 10 tests.

Our approach is therefore to flip some small percentage of all true hypothesis vector elements (5-20%). This approach also can not guarantee perfect accuracy. The reason is that we might have failed to randomly flip a specific test choice for a true hypothesis in the way that corresponds to an actual noisy choice a person made.

The extent to which the wrong hypotheses are selected, averaged across many simulated designs, will be seen in ground-truth simulations below (Figures 3.1 and 3.2). The asymptotic probability of choosing the correct hypothesis is about 90% for the reasonable noise rates we use. That is, even if all tests are conducted, the fact that noise is injected randomly means we sometimes choose a wrong hypothesis (if the bit flipping just happened to match choices). This is an unavoidable price for a realistic noise model such as this one, which allows for human error to lead to model indistinguishability in some cases. Next we describe the method formally; The notation is adopted from Computer Science (Golovin and Krause (2010c)).

3.2.3 Bayesian Active Learning in the Noiseless Case

In the Bayesian active learning problem, we would like to distinguish among a given set of hypotheses $\mathcal{H} = \{h_1, \ldots, h_n\}$ by performing tests from a set $\mathcal{T} = \{1, \ldots, N\}$ of possible tests. Running test t incurs a cost of $c(t)$ and produces an outcome from a finite set of outcomes $\mathcal{X} = \{1, 2, \ldots, \ell\}$.

We let H denote the random variable which equals the *true* hypothesis, and model the outcome of each test t by a random variable X_t taking values in $\mathcal{X}$. We denote the observed outcome of test t by x_t. We further suppose we have a prior distribution P modeling our assumptions on the joint probability $P(H, X_1, \ldots, X_N)$ over the hypotheses and test outcomes.

In the noiseless case, we assume that the outcome of each test is deterministic given the true hypothesis, i.e., for each $h \in \mathcal{H}$, $P(X_1, \ldots, X_N \mid H = h)$ is a deterministic distribution. Thus, each hypothesis h is associated with a particular vector of test outcomes. We assume, w.l.o.g., that no two hypotheses lead to the same outcomes for all tests. Thus, if we perform all tests, we can uniquely determine the true hypothesis. Hereby, $\mathbf{x}_\mathcal{A} \in \mathcal{X}^\mathcal{A}$ is a vector of outcomes indexed by a set of tests $\mathcal{A} \subseteq \mathcal{T}$ that we have performed so far. After having made observations $\mathbf{x}_\mathcal{A}$, we can rule out inconsistent hypotheses. We denote the set of hypotheses consistent with event Λ (often called the *version space* associated with Λ) by $(\Lambda) := \{h \in \mathcal{H} : P(h \mid \Lambda) > 0\}$.

However in most applications we will wish to avoid performing every possible test, as this is prohibitively

expensive. Our goal is to find an adaptive policy for running tests that allows us to determine the value of H while minimizing the cost of the tests performed. Formally, a policy π (also called a conditional plan) is a partial mapping π from partial observation vectors $\mathbf{x}_\mathcal{A}$ to tests, specifying which test to run next (or whether we should stop testing) for any observation vector $\mathbf{x}_\mathcal{A}$.

We can define the expected cost of a policy π by

$$c(\pi) := \sum_h P(h)c(\mathcal{T}(\pi, h))$$

where $\mathcal{T}(\pi, h) \subseteq \mathcal{T}$ is the set of tests run by policy π in case $H = h$.

Our goal is to find a feasible policy π^* of minimum expected cost, i.e.,

$$\pi^* = \arg\min \{c(\pi) : \pi \text{ is feasible}\} \tag{3.2.1}$$

We call a policy *feasible* if it is guaranteed to uniquely determine the correct hypothesis. That is, upon termination with observation $\mathbf{x}_\mathcal{A}$, it must hold that $|(\mathbf{x}_\mathcal{A})| = 1$.

A policy π can be naturally represented as a decision tree T^π, and thus problem (3.2.1) is often called the *Optimal Decision Tree* (ODT) problem. Unfortunately, obtaining even an approximate policy π that is near-optimal[10] is NP-hard (Chakaravarthy *et al.* (2007)).

Hence, various heuristics are employed to solve the Optimal Decision Tree problem and its variants. Two of the most popular heuristics are to select tests greedily to maximize the *information gain* (IG) conditioned on previous test outcomes, and *generalized binary search* (GBS).

Both heuristics are greedy (myopic), and after having made observations $\mathbf{x}_\mathcal{A}$ will select

$$t^* = \arg\max_{t \in \mathcal{T}} \Delta_{\text{Alg}}(t \,|\, \mathbf{x}_\mathcal{A}) / c(t),$$

where $\text{Alg} \in \{\text{IG}, \text{GBS}\}$.

Here,

$$\Delta_{\text{IG}}(t \,|\, \mathbf{x}_\mathcal{A}) := \mathbb{H}(\mathbf{X}_\mathcal{T} \mid \mathbf{x}_\mathcal{A}) - \mathbb{E}_{x_t \sim X_t | \mathbf{x}_\mathcal{A}}[\mathbb{H}(\mathbf{X}_\mathcal{T} | \mathbf{x}_\mathcal{A}, x_t)] \tag{3.2.2}$$

is the marginal information gain measured with respect to the Shannon entropy $\mathbb{H}(\mathbf{X}) := \mathbb{E}_\mathbf{x}[-\log_2 P(\mathbf{x})]$, and

$$\Delta_{\text{GBS}}(t \,|\, \mathbf{x}_\mathcal{A}) := P((\mathbf{x}_\mathcal{A})) - \sum_{x \in \mathcal{X}} P(X_t = x \mid \mathbf{x}_\mathcal{A}) P((\mathbf{x}_\mathcal{A}, X_t = x)) \tag{3.2.3}$$

is the expected reduction in version space probability mass. Thus, both heuristics greedily (or myopically)

[10] For which the cost $c(\pi) \leq c(\pi^*) \cdot o(\log(n))$

chooses the test that maximizes the benefit-cost ratio, measured with respect to their particular benefit functions[11].

Interestingly, despite its myopic nature GBS has been shown(Kosaraju *et al.* (1999), Dasgupta (2004), Guillory and Bilmes (2009), Golovin and Krause (2010c)) to obtain near-optimal expected cost. The strongest known bound is

$$c(\pi_{GBS}) \leq c(\pi^*) \left(\ln(1/p_{\min}) + 1\right) \tag{3.2.4}$$

where $p_{\min} := \min_{h \in \mathcal{H}} P(h)$.

The result above is proved by exploiting the fact that f_{GBS} is *adaptive submodular* and *strongly adaptively monotone*(Golovin and Krause (2010c))[12]. In Appendix **??**, the concept of adaptive submodularity to used to prove approximation guarantees for Bayesian active learning with noisy observations.

3.2.4 The Equivalence Class Determination Problem

Generally, we introduce the *Equivalence Class Determination* problem[13], where our set of hypotheses $\mathcal{H}$ is partitioned into a set of m equivalence classes $\{\mathcal{H}_1, \ldots, \mathcal{H}_m\}$ so that $\mathcal{H} = \biguplus_{i=1}^{m} \mathcal{H}_i$, and the goal is to determine which class $\mathcal{H}_i$ the true hypothesis lies in.

We now wish to consider the Bayesian active learning problem where tests can have noisy outcomes. Our general strategy is to reduce the problem of noisy observations to the noiseless setting. To gain intuition, consider a simple model where tests have binary outcomes, and we know that the outcome of exactly one test, chosen uniformly at random unbeknown to us, is flipped. If any pair of hypotheses $h \neq h'$ differs by the outcome of at least three tests, we can still uniquely determine the correct hypothesis after running all tests. In this case we can reduce the noisy active learning problem to the noiseless setting by, for each hypothesis, creating N "noisy" copies, each obtained by flipping the outcome of one of the N tests. The modified prior P' would then assign mass $P'(h') = P(h)/N$ to each noisy copy h' of h. The conditional distribution $P'(\mathbf{X}_{\mathcal{T}} \mid h')$ is still deterministic (obtained by flipping the outcome of one of the tests). Thus, each hypothesis h_i in the original problem is now associated with a set $\mathcal{H}_i$ of hypotheses in the modified problem instance. However, instead of selecting tests to determine which noisy copy has been realized, we only care which set $\mathcal{H}_i$ is realized.

The ECD algorithm. We define a set of edges $\mathcal{E} = \cup_{1 \leq i < j \leq m} \{\{h, h'\} : h \in \mathcal{H}_i, h' \in \mathcal{H}_j\}$, consisting of all (unordered) pairs of hypotheses belonging to distinct classes. These are the edges that must be

[11] It turns out that for the (noiseless) Optimal Decision Tree problem, these two heuristics are equivalent (Zheng *et al.* (2005))

[12] Since $f_{GBS}(S, h) := 1 - P((\mathbf{x}_S(h))) + P(h)$

[13] Bellala et al. simultaneously studied ECD Bellala *et al.* (2010), and used it to model active learning with noise Bellala *et al.* (2009). They developed an extension of GBS for ECD.

cut, by which we mean for any edge $\{h, h'\} \in \mathcal{E}$, at least one hypothesis in $\{h, h'\}$ must be ruled out (i.e., eliminated from the version space). Hence, a test t run under true hypothesis h is said to cut edges $\mathcal{E}_t(h) := \{\{h', h''\} : h'(t) \neq h(t) \text{ or } h''(t) \neq h(t)\}$. See Fig. 3.1 for an illustration. We define a weight function $w : \mathcal{E} \to \mathbb{R}_{\geq 0}$ by $w(\{h, h'\}) := P(h) \cdot P(h')$. We extend the weight function to an additive (modular) function on sets of edges in the natural manner, i.e., $w(\mathcal{E}') := \sum_{e \in \mathcal{E}'} w(e)$. The objective f_{EC} that we will greedily maximize is then defined as the weight of the edges cut (EC):

$$f_{EC}(\mathcal{A}, h) := w\Big(\bigcup_{t \in \mathcal{A}} \mathcal{E}_t(h)\Big) \tag{3.2.5}$$

The objective f_{EC} is strongly adaptively monotone and adaptively submodular, which leads to proofs of theoretical guarantees using the results of Golovin and Krause (2010c). Based on the objective f_{EC}, we can calculate the marginal benefits for test t upon observations $\mathbf{x}_\mathcal{A}$ as

$$\Delta_{\text{EC}}(t \,|\, \mathbf{x}_\mathcal{A}) := \mathbb{E}_H[f_{EC}(\mathcal{A} \cup \{t\}, H) - f_{EC}(\mathcal{A}, H) \;|\; \mathbf{x}_\mathcal{A}]. \tag{3.2.6}$$

We call the adaptive policy π_{EC} that, after observing $\mathbf{x}_\mathcal{A}$, greedily selects test

$$t^* \in \arg\max_t \Delta_{\text{EC}}(t \,|\, \mathbf{x}_\mathcal{A}) / c(t), \tag{3.2.7}$$

the EC^2 algorithm (for *equivalence class edge cutting*).

3.2.5 Noisy Observations

A major challenge when dealing with noisy observations is that it is not always possible to distinguish distinct hypotheses. Even after we have run all tests, there will generally still be uncertainty about the true hypothesis, i.e., the posterior distribution $P(H \mid \mathbf{x}_\mathcal{T})$ obtained using Bayes' rule may still assign non-zero probability to more than one hypothesis. If so, uniquely determining the true hypothesis is not possible. Instead, we imagine that there is a set $\mathcal{D}$ of possible *decisions* we may make after (adaptively) selecting a set of tests to perform and we must choose one (e.g., we must decide how to treat the medical patient, which scientific theory to adopt, or which classifier to use, given our observations). Thus our goal is to gather data to make effective decisions (Howard (1966)). Formally, for any decision $d \in \mathcal{D}$ we take, and each realized hypothesis h, we incur some loss $\ell(d, h)$. Decision theory recommends, after observing $\mathbf{x}_\mathcal{A}$, to choose the decision d^* that minimizes the *risk*, i.e., the expected loss, namely

$$d^* \in \arg\min_d \mathbb{E}_H[\ell(d, H) \mid \mathbf{x}_\mathcal{A}]. \tag{3.2.8}$$

A natural goal in Bayesian active learning is thus to adaptively pick observations, until we are guaranteed to make the same decision (and thus incur the same expected loss) that we would have made had we run *all* tests. Thus, we can reduce the noisy Bayesian active learning problem to the ECD problem by defining the equivalence classes over all test outcomes that lead to the same minimum risk decision. Hence, for each decision $d \in \mathcal{D}$, we define

$$\mathcal{H}_d := \{\mathbf{x}_\mathcal{T} : d = \arg\min_{d'} \mathbb{E}_H[\ell(d', H) \mid \mathbf{x}_\mathcal{T}]\}. \tag{3.2.9}$$

If multiple decisions minimize the risk for a particular $\mathbf{x}_\mathcal{T}$, we break ties arbitrarily. Identifying the best decision $d \in \mathcal{D}$ then amounts to identifying which equivalence class $\mathcal{H}_d$ contains the realized vector of outcomes, which is an instance of ECD. More generally, we have the following result for EC^2.

Theorem 9. *Fix hypotheses $\mathcal{H}$, tests $\mathcal{T}$ with costs $c(t)$ and outcomes in $\mathcal{X}$, decision set $\mathcal{D}$, and loss function ℓ. Fix a prior $P(H, \Theta)$ and a function $\mathbf{x}_\mathcal{T} : \mathcal{H} \times \text{supp}(\Theta) \to \mathcal{X}^N$ which define the probabilistic noise model. Let $c(\pi)$ denote the expected cost of π incurs to find the best decision, i.e., to identify which equivalence class $\mathcal{H}_d$ the outcome vector $\mathbf{x}_\mathcal{T}$ belongs to. Let π^* denote the policy minimizing $c(\cdot)$, and let π_{EC} denote the adaptive policy implemented by* EC^2. *Then it holds that*

$$c(\pi_{EC}) \leq \left(2 \ln \left(\frac{1}{p'_{\min}}\right) + 1\right) c(\pi^*),$$

where $p'_{\min} := \min_{h \in \mathcal{H}} \{P(h, \theta) \ : \ P(h, \theta) > 0\}$.

Stated simplistically, for economic experiments, where each test has the same cost, the policy implemented by EC^2 will, in the *worst-case*, require $2 \ln(1/p_{min})$ more number of tests than the *optimal* policy (which is infeasible to compute).

3.3 Implementation

The goal is to adaptively select a sequence of tests to present to a subject in order to distinguish which of the five models best explains the subject's responses. Here a test is a pair of lotteries, (L_1, L_2). Based on the theory that represents behaviour, one of the lotteries would be preferred to the other, denoted by a binary response $y \in \{1, 2\}$.

Each lottery consists of an amount $x \in \{\$5, \$10, \cdots, \$50\}$ and a delay t from 0 (payable immediately) to 90 days. By considering all non-identical pairs of lotteries, we obtained the set of possible tests (roughly 14,000).

The BROAD algorithm and effective objective criteria, alongwith an accelerated greedy version is outlined in Chapter 2. Here we use simulations in section 3.3.2 as a tool for analyzing the efficiency of our experimental setup. The accuracy of the method in finding the true model, with the number of tests reveals that we can expect reasonable answers within 50 questions. Furthermore the Equivalence class edge cutting criteria is demonstrated to be superior to information gain with error in responses.

3.3.1 Experimental Setup

Subjects were recruited from the UCLA and Caltech student population. Experimental conditions and payment procedure can possibly influence subject's time preference behaviour. We choose a standard paradigm of presenting two choices, one with a smaller amount with a smaller delay, and a larger amount with greater delay; the details are provided in section 3.3.1. Specifying the delay as calendar time, instead of delay in time (number of days from now) has been shown to increase patience, Read *et al.* (2005). Here we present delay as time, in accord with earlier studies that have compared time preference models, Benhabib *et al.* (2010). To avoid any transaction costs for immediate versus delayed payments, the same method (online payments) is used when the chosen delay is 0.

To conduct the experiment, we used a text display that showed the two lotteries. The instructions to the subjects are provided in Appendix A.

3.3.2 Ground Truth Analysis

We compared the Information Gain criteria with Equivalence Class Edge-Cutting and a method that picked tests randomly from the set of possible tests.

We evaluated the ability of the algorithm to recover the true model based on simulated responses. The responses were generated using a trembling hand function, with the probability of response $y = 1$ given by

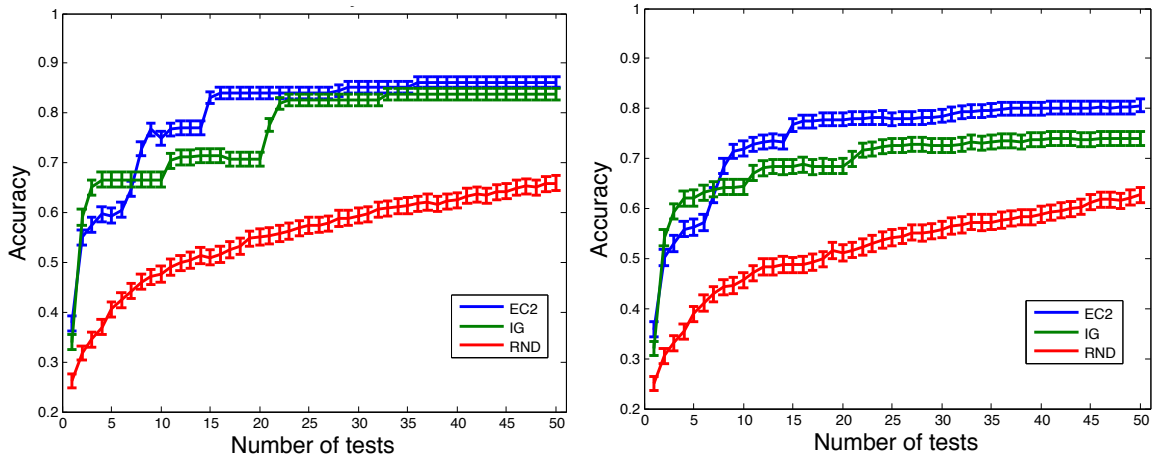


Table 3.1: Performance of the algorithms when there is no error, $\epsilon = 0.0$ (left) and when probability of error, $\epsilon = 0.05$ (right).

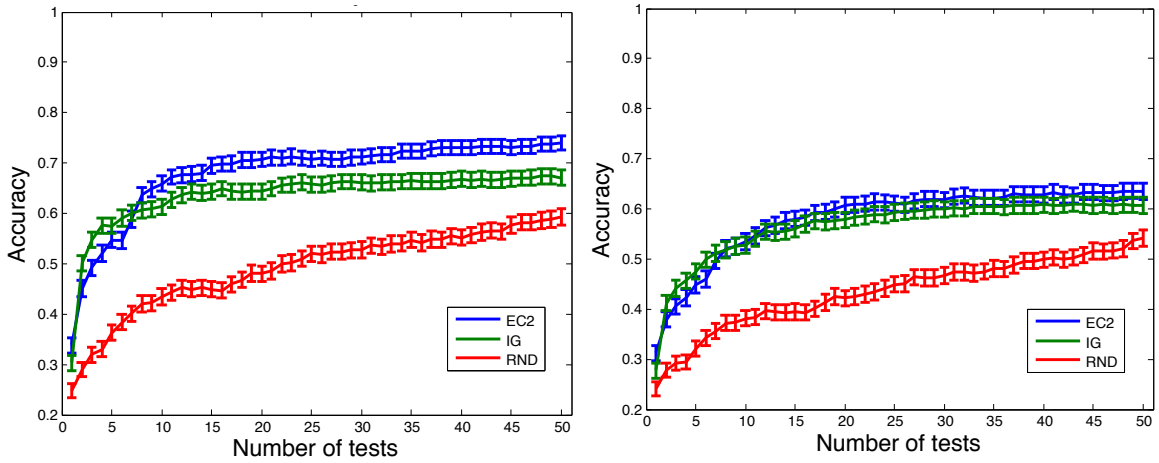


Table 3.2: Performance of the algorithms when probability of error, $\epsilon = 0.10$ (left) and $\epsilon = 0.20$ (right).

$P(y = 1) = 1 - \epsilon$, with the erroneous response $y = 2$ occurring with probability ϵ.

We evaluated the speed and ability of the three methods to recover the true model based on a large sample of simulated response sequences. In each sequence, a specific true model is assumed, which generates simulated responses to the sequence of test choices generated by each method. The conclusion the method draws about which model is likely to have produced the observed responses is then compared to the true (known) model. This model-recovery exercise is sometimes called "ground truth" analysis. The responses were generated using a trembling hand function, with a known probability ϵ of making the wrong choice due to noise.

We chose discrete prior parameter values for different theories based on our experimental setup and, more generally, on previous estimates reported in literature. Each parameter has 11 different values, equally spaced between the lowest to the highest. This results in 11 different parameter values for the one-parameter models, and 121 values for the two-parameter models.

- The Exponential discounting model (Exp) has one discounting parameter, $r \in \{0.0, \cdots, 0.02\}$. For the maximum delay of 90 days, the highest discount rate $r = 0.02$, results in a discount factor of 0.165. The decision-maker is indifferent between \$8.26 and \$50.00 after 90 days.

- The Hyperbolic discounting model (Hyp) also has one discounting parameter, $r \in \{0.0, \cdots, 0.2\}$. For the highest discount rate of $r = 0.2$, the model is indifferent between \$2.63 now and \$50.00 after 90 days.

- The Quasi-hyperbolic discounting model (Quasi-Hyp) has 2 discounting parameters: an exponential discounting factor r and α. We use the same range of values for r as in the exponential model ($r \in \{0.0, \cdots, 0.02\}$); And $\alpha \in \{0.5 \cdots 1.0\}$. For the highest discount rates, $r = 0.2$ and $\alpha = 0.5$, the model is indifferent between \$1.31 now and \$50 after 90 days. In Albrecht *et al.* (2011), the median α from individual experimental data was 0.86, with 24 of 27 participants having $\alpha < 1$.

- The Fixed-cost discounting model (Fixed) has 2 discounting parameters: an exponential discounting factor has range $r \in \{0.0, \cdots, 0.02\}$ a fixed time cost $b \in \{0 \cdots 20\}$. Benhabib *et al.* (2010) found a present bias / present premium of about \$4 among amounts of \$10 through \$100.

- Finally, the Generalized-hyperbolic discounting model (GenHyp) has 2 discounting parameters: θ and α. We chose parameter values to match the range of discount factors generated by the previous 4 models. For θ we chose $\theta \in \{0.02 \cdots 0.22\}$; and we chose $\alpha \in \{0.5 \cdots 1.0\}$. For the higest discount rates, $\theta = 0.22$ and $\alpha = 1$, the model is indifferent between \$6.87 now and \$50.00 after 90 days. When the ratio $\theta/\alpha < 1$, we get more hyperbolic discounting, wheareas when the ratio is large, the model begins to approximate hyperbolic discounting. Typical values that are reported in literature for θ/α are in the range $[0.45, 0.78]$ (McKerchar *et al.* (2010, 2009)).

For the ground truth analysis, each of the five model classes started with an equal overall prior belief of $1/5$, spread uniformly across all parameter value configurations in that model class. (For example, each of the 121 specific parameter configurations for GenHyp was assigned prior $1/5$ divided by 121.) In each ground truth run, a model and parameter configuration was chosen at random. Fifty test choices were then generated using each of the five methods. In each such run, we record whether each of the five methods is accurate after each question. Accuracy in a run is defined as assigning the highest posterior probability to the ground-truth model (for the overall model in the same class) in that run. The procedure was repeated for 1000 runs.

For the ground truth analysis, we generated simulated responses by first choosing a model at random, and then the parameter values for the model at random. We repeated this for a 1000 trails.

Figures 3.1 and 3.2 show the performance of the 3 methods, in terms of the accuracy of recovering the true model with the number of tests, for different error rates ($\epsilon = 0.0, 0.05, 0.10, 0.20$). We find that EC2 outperforms InfoGain significantly, which outperforms Random for all noise levels.

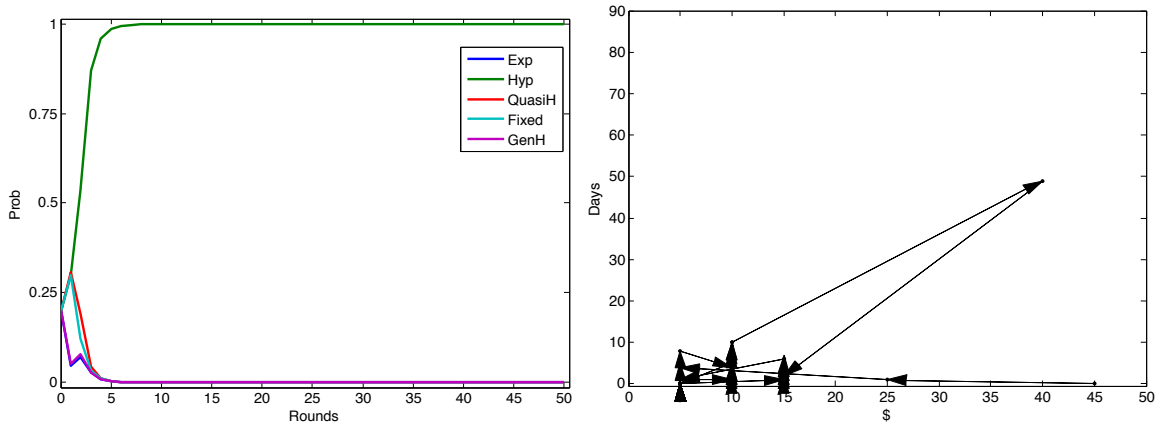


Table 3.3: The posterior probability of the models over 50 rounds for a ”Hyperbolic” type (left) and the choice path: amount and delay of the options selected (right).

3.4 Analysis

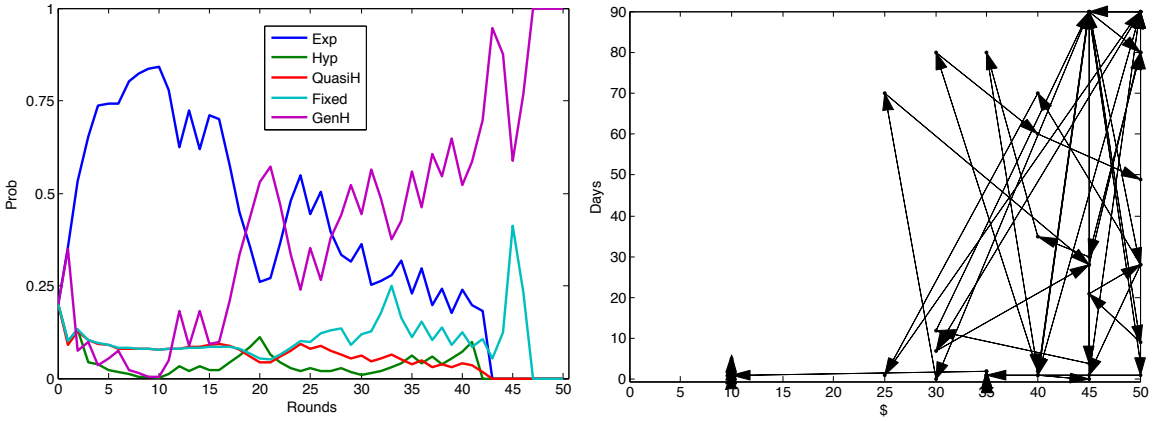


Table 3.4: The posterior probability of the models over 50 rounds for a ”General-Hyperbolic” type (left) and the choice path: amount and delay of the options selected (right).

We tested 40 subjects in 3 separate sessions at Caltech (n=10) and UCLA to determine which model fit their behaviour best. We followed the IRB protocols of Caltech and UCLA for the conduct of human subject experiments. Subjects were presented 50 tests using our algorithm. To incentivize the subjects, one of these tests was picked at random, and subjects received payment based the outcome of their chosen lottery.

The delayed payments were made into the debit cards for subjects at UCLA, and as Amazon gift cards for subjects at Caltech. Both forms of deferred payment can be used to purchase a similar variety of goods as cash. To prevent any transaction effects, the immediate payment (if delay = 0 were chosen) was also made using the same method and sent immediately after the experiment.

We observe heterogeneity in the population. We illustrate with two examples from actual subject behaviour.

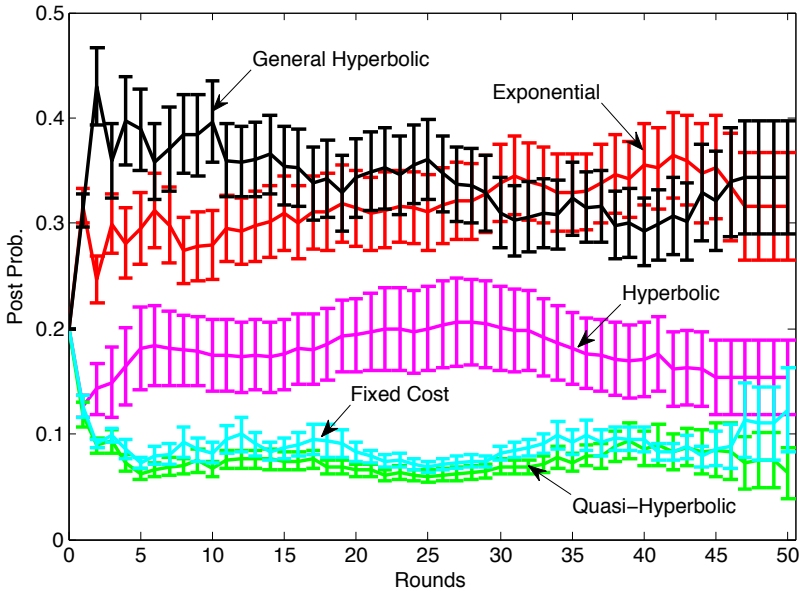


Figure 3.2: The posterior distribution of the 5 models after each round. The aggregate mean and standard error is shown for 40 subjects.

In Figure 3.3, we plot the posterior probability with the number of rounds for a subject that was quickly classified as a Hyperbolic type (with 5 rounds). The also plot the "Choice Path" over the 50 rounds: the amount and delay of the selected options. We find that the subject options that had low delay for various amounts. This illustrates the algorithm's ability to narrow down the parameters that best describe the subject's preference and offer choices that best distinguish the theories. In figure 3.4, we plot the posterior probability and choice path of a subject whose behaviour is eventually captured by the General-Hyperbolic model.

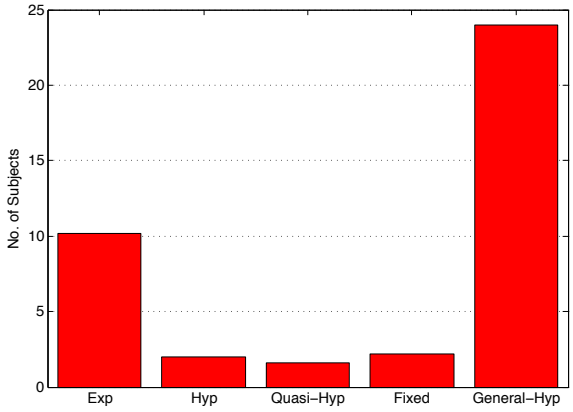


Figure 3.3: The classification of 40 subjects to each of the 5 models.

Figure 3.4 shows the posterior distribution of the models after each round, aggregate over all subjects. It is clear that the posterior distribution of the generalized hyperbolic discounting and exponential discounting

functions were significantly higher than the other 3 models. The evidence for the "present-bias" models, as well as the hyperbolic discounting model, was low.

In figure 3.4 we show a histogram of the number of subjects classified according to each model. After each round, we obtain a posterior distribution over the 5 models. Behaviour can be classified into one of the models, by taking the maximum-a-posteriori estimate. To ensure robustness, we take the average MAP estimate of the last 5 rounds.

3.4.1 Hyper-patience

13 subjects in our experiments always chose the greater amount irrespective of the delay. These "hyper-patient" subjects' behaviour was best captured by the Generalized-Hyperbolic model. We wanted to rule out the possibility that this classified was not due to the parameter support we had used for the models.

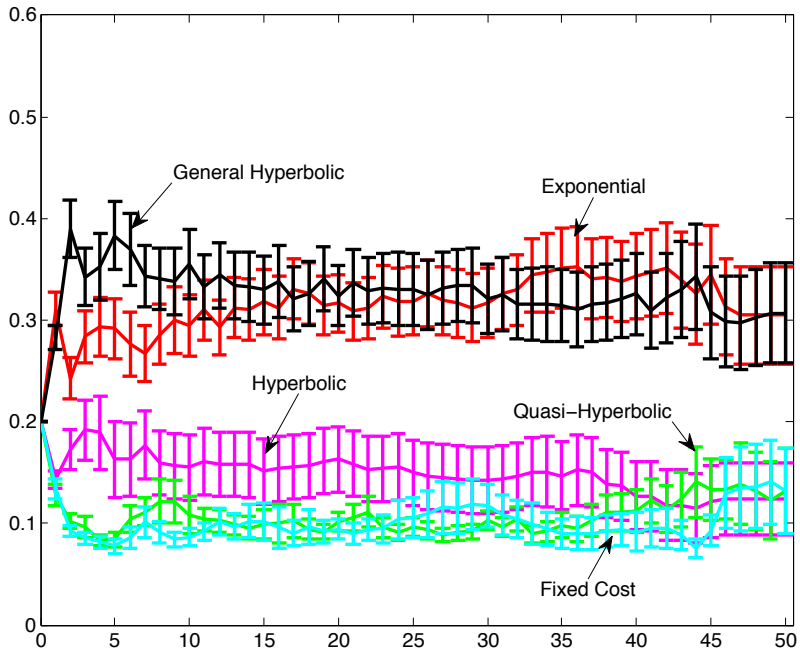


Figure 3.4: The posterior distribution of the 5 models after each round. The aggregate mean and standard error is shown for 40 subjects.

It is conventional, in the literature, to set the discount factor to positive in the exponential model. If we allow the discount factor to take small negative values, then the model would encode "hyper-patience" a preference for the larger amount, irrespective of delay.

The changed the parameter support for all 5 models, so that the models exhibit a preference for the larger amount irrespective of delay. This meant, adding support for small negative values for some of the parameters. The grid size increased from 11, to 12, equally-spaced values. Specifically, the parameter values ranged as

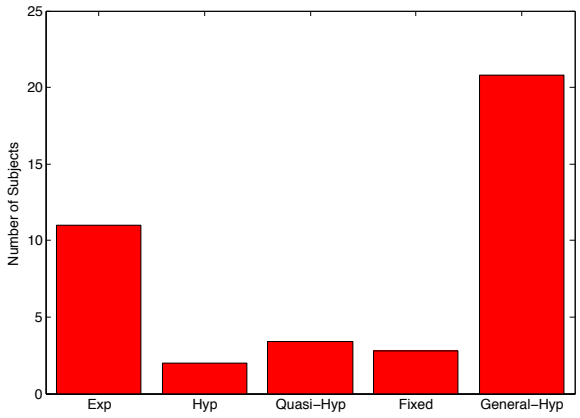


Figure 3.5: The classification of 40 subjects to each of the 5 models with expanded parameter support for "hyper-patience".

follows:

◦ For Exponential discounting model, the support of r included the negative value of $r = -0.002$. Thus, $r \in \{-0.002 \cdots 0.02\}$.

◦ For Hyperbolic disocounting model, support of r included a negative value of $r = -0.02$. Thus, $r \in \{-0.02 \cdots 0.2\}$.

◦ For Quasi-Hyperbolic discounting model, $r \in \{-0.002 \cdots 0.02\}$, and values of α were the same; $\alpha \in \{0.5 \cdots 1.0\}$.

◦ For Fixed-cost discounting model, $r \in \{-0.002 \cdots 0.02\}$, and values of b were the same; $b \in \{0 \cdots 20\}$.

◦ For Generalized-Hyperbolic discounting model, support of θ included a negative value $\theta = -0.02$. Thus $\theta \in \{-0.02 \cdots 0.22\}$, and values of α were the same; $alpha \in \{0 \cdots 20\}$.

We re-analyzed the choice data from our experiment, by running the algorithm with the choices that were made by the subjects. The aggregate posterior probability and classification is shown in figures 3.4.1 and 3.4.1 respectively. The results are statistically not different from above: the Generalized-Hyperbolic model has the highest number of classifications, followed by the exponential model. Support for "present-bias" models was weak.

3.4.2 Distribution of Parameters

At the end of the experiment, we obtain a posterior distribution over all hypotheses (theory and specific parameters). We sum over all parameters within a theory to obtain the posterior distribution over the theories, and obtain the MAP classification of the subjects. Given the final classification to a theory, we can also obtain the MAP estimate of specific parameters. The MAP parameter estimates for the theories were:

- Exponential discounting (Exp) had $n = 11$ classifications; average values of parameters (with standard errors) were: $r = 0.002(\pm 0.003)$.

- Hyperbolic discounting (Hyp) had $n = 2$ classifications; average values of parameters (with standard errors) were: $r = 0.010(\pm 0.028)$.

- Quasi-Hyperbolic discounting (Quasi-Hyp) had $n = 4$ classifications; average values of parameters (with standard errors) were: $r = 0.003(\pm 0.008), \alpha = 0.68(\pm 0.01)$.

- Fixed Cost discounting (Fixed) had $n = 3$ classifications; average values of parameters (with standard errors) were: $r = 0.002(\pm 0.002), b = 7.0(\pm 1.0)$.

- Generalized-Hyperbolic discounting (General-Hyp) had $n = 20$ classifications; average values of parameters (with standard errors) were: $\theta = 0.032(\pm 0.008), \alpha = 0.84(\pm 0.17)$.

Since the objective for EC^2 is to distinguish between theories, the experimental design and the sequence of questions asked is not optimized towards increasing the resolution of the parameter estimates, or even narrowing in on the correct parameters. Our methodology can easily be extended to estimating the parameters of a specific theory, such as prospect theory, with a high resolution.

3.5 Discussion

We applied Bayesian Rapid Optimal Adaptive Design (BROAD), a recent methodological advance in experimental testing (ray2013), to compare 5 different types of time preference models. We found the strongest evidence for generalized-hyperbolic discounting, and some evidence for exponential discounting. Recent studies have reported evidence for "present bias" models, for which our experimental results show weak evidence. Our payment method eliminated any transaction costs for future payments, since immediate payments were also delivered through the same channel, so there is no "present bias" by design.

There are several avenues for empirical research that can be motivated by our study. While generalized-hyperbolic discounting generalizes exponential and hyperbolic models, Benhabib *et al.* (2004) presented a generalized model that nests all of the models that we compared. Although their model is expressive enough, Benhabib *et al.* (2004)'s particular data had little power to distinguish between the fits of different functional forms. This is a challenge for models with multiple parameters, in that a large number of data points will be required to obtain statistical power to distinguish the functional forms of different discounting models. BROAD economizes the number of samples required to distinguish among models, and can be used to compare a very generalized model with models that are more parsimonious in the number of parameters.

The class of models compared in this study, uses a subjective perception of money (as a function of time), while the perception of time is linear. This class of models are common amongst economists. Another class of models, with roots in biology and psychology, models the subjective perception of time which leads to inconsistencies in inter-temporal choices, e.g. Ebert and Prelec (2007); Bleichrodt *et al.* (2009); Ray and Bossaerts (2011). BROAD can be used to compare these essentially different classes of models.

Finally, time preference behaviour depends on domain and context. Preferences can change based on the type of reward: monetary or consumables (Estle *et al.* (2007)), and can be attributed to pathological behaviours, such as addiction and gambling (Madden and Bickel (2010)). BROAD methods can be utilized to rapidly evaluate models in different contexts and with different subject populations, especially subjects that are limited in attention (pathological behaviours, children and animals).

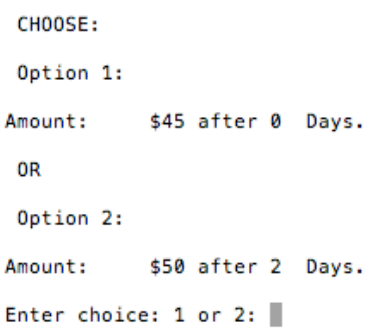
3.6 Instructions

This is an experiment on the economics of decision-making. If you follow the instructions and make good decisions you can earn a considerable amount of money, which will be paid to you in the form of Amazon Gift cards at a future date, as determined by your choices. The funds have been supplied by a foundation that supports this research.

The experiment has 50 rounds. In each round you will choose between one of two options: Each option has an amount, $X, payable after T days.

The amounts will range from $5 up to $50, and the days will range from 0 (given immediately) up to 90 days (3 months later).

For example, you might see the following screen:



In option 1, you will get $45 by today. In option 2, you will get $50 2 days from today.

You must enter your choice into the textbox by pressing 1 or 2, followed by the Enter key.

There is no time limit for viewing the graph and entering your choice. After you enter your choice, the screen will say REST. Another choice between two lottery options will be displayed within a couple of seconds.

After finishing all 50 rounds, we will randomly pick one of the 50 pairs of options you were asked about to determine your additional earnings. Each option pair is equally likely to be chosen so you should treat each one as if it could be the paid round.

Depending on your choice, the Gift card for the amount $X will be sent to you by email after T days (or immediately if T = 0 days). For example, if you chose option 2 in the case above, your amount will be emailed to you 2 days from today.

3.7 Subjective Time

Suppose we are asked to choose between \$10 now or \$11 tomorrow. We may prefer the \$10 immediately rather than the \$11 received after a day. However, if we are offered to choose between \$10 to be received after 364 days, or \$11 after 365 days, we often prefer to wait the additional day for the extra dollar. After waiting 365 days, the latter choice becomes one between an immediate \$10 or \$11 tomorrow. Now we would want to reverse it, asking for the \$10 immediately, rather than waiting the extra day we seemed to have been willing to accept in the past. This time inconsistency can be modeled using hyperbolic discounting of future rewards Laibson (1997). To avoid inconsistencies, rewards should be discounted exponentially over time Sutton and Barto (1998).

Here, we conjecture that hyperbolic discounting has a rational explanation, based on the generally accepted principal that the animal and human biological clocks tick at a different rate from the calendar clock. Animals and humans are indeed known to maintain an internal representation of time that differs from standard calendar time and whose properties change with time scale, from microseconds up to years, with representations of (calendar) time at larger scales showing the highest variability Buonomano (2007). In humans, drugs Meck (1996); Wittmann *et al.* (2007), and age Mischel *et al.* (1992), among other things, are known to influence time perception. The neurobiological mechanism behind the biological clock has recently become a topic of intense study Buhusi and Meck (2005).

We start from the preposition that discounting is exponential, as required for time-consistent choice. At the same time, we posit that the internal representation of time (under which discounting occurs) varies stochastically (randomly) from calendar time. This is illustrated in Figure 3.6: two events at times 0 and Δ in calendar are experienced to occur at $t_0 = F(0)$ and $t_\Delta = F(\Delta)$ in biological time, where F is some stochastic (random) transformation F.

Now consider two events at a later time s and $s + \Delta$, separated by the same amount of (calendar-time) delay Δ. The times s and $s + \Delta$ are transformed to $t_s = F(s)$ and $t_{s+\Delta} = F(s + \Delta)$ in biological time. Although the delay between the pair of events is the same in calendar time, the corresponding delay between the events in biological time, $t_\Delta - t_0$ and $t_{s+\Delta} - t_s$ are generally different.

We assume that the biological clock is positive quadrant dependent Esary *et al.* (1967). Intuitively, this means that if t_s is small, the chance increases that the subsequent interval $t_\Delta - t_s$ is small too. It implies positive autocorrelation, i.e., $cov(t_s, t_{s+\Delta} - t_s) > 0$. And crucially, it implies that discount factors are positive autocorrelated: $cov(\exp(-t_s), \exp(-(t_{s+\Delta} - t_s))) > 0$. We also assume that increments in biological time are stationary: the distribution of time changes does not depend on when they occur; the distribution of $t_\Delta - t_0$ is the same as that of $t_{s+\Delta} - t_s$. The following theorem states the main result.

Theorem: Positive dependence in biological time implies temporal choice inconsistency The rigorous proof of the theorem is provided in the Appendix, that the reader is encouraged to peruse. Here we aim to provide the intuition with a simple numerical example. We envisage delivery of one dollar now or K dollars at Δ (a point in calendar time). K is to be chosen so that there is indifference. We then compare delivery of one dollar at $s = 2$ with delivery of K dollars at $s = 2 + \Delta$.

The biological clock is as follows. Either s feels like it takes 1 unit of (biological) time, or it takes 3 units, with an average of 2 units. Δ is on average half the length; it feels either short (0.5 units) or long (1.5), with an average of 1 unit.

The biological clock is positive quadrant dependent. We will make the dependence perfect, to simplify the argument. When s feels like it takes 1 unit, the subsequent Δ will take 0.5 unit; when s feels long (2 units), the subsequent Δ is long as well, at 1.5 units.

As mentioned, K is set so that there is indifference between getting one dollar now and K dollars at Δ. Today's value of K is $0.5 * exp(-0.5) + 0.5 * exp(-1.5)) * K$, because with fifty percent probability, delivery is felt like taking place in 0.5 units of (biological) time, and with 50% chance, it feels like it takes place 1.5 time units in the future. We set K so that today's value equals one dollar. So, $K = 2/(exp(-0.5) + exp(-1.5)) = 2.411$.

Essentially, K is set so that the gain of having to wait a short time (only 0.5 units of biological time) is offset by the loss in value for having to wait a long time (1.5 units). The former gain (relative to today's one dollar) is $2.411 * exp(-0.5) - 1 = 0.462$, or 46.2%; the latter loss equals $1 - 2.411 * exp(-1.5) = (1 - 0.538) = 46.2\%$. The gains and losses offset.

As for delivery of K at $s + \Delta$, notice that, while the gain and loss from waiting an extra Δ beyond s have equal probability of occurring, *they are discounted differently*. The gain occurs when $s = 2$ arrives early under the biological clock (1 unit); this is weighted more heavily, because it is discounted with only $exp(-1)$; its weighted value is $exp(-1) * 1.462 = 0.538$. The loss when $s = 2$ arrives late (3 units under the biological clock) is weighted less, because it occurs farther in time; it is discounted with $exp(-3)$; the weighted value in that case is $exp(-3) * 0.538 = 0.027$. In expectation, the value of receiving K at $s + \Delta$ equals $(0.538 + 0.027)/2 = 0.283$.

Compare this to getting one dollar at $s = 2$. There is no loss or gain (one always gets one dollar). With 50% probability, the dollar arrives early (1 unit in biological time), and its discounted value is $exp(-1) = 0.368$, and with 50% chance the dollar arrives late (3 units under the biological clock), and its discounted value is $exp(-3) = 0.050$. In expectation, the value equals $(0.368 + 0.050)/2 = 0.209$. This is strictly less than the

value of getting K delayed (0.283).

Hence, while K was set to be indifferent between receiving one dollar now and K dollars at Δ, the promise of K at $s + \Delta$ is worth more than getting one dollar at s.

The astute reader will have noticed that the positive dependence of the biological clock is not really needed to get time inconsistencies. They occur also with negative dependence. However, with positive dependence, the decision maker will always prefer to *wait* when comparing options in the future for which she is indifferent now. That is, she looks more patient when deciding about payoffs in the future. This is the classical time inconsistency that has led to modeling of intertemporal choice using hyperbolic discounting. With negative dependence, the decision maker looks more patient in the immediate future than when considering options in the far future.

3.7.1 Simulations

We now illustrate that hyperbolic discounting provides a good fit to the choices resulting from positive temporal dependence of biological time. To model biological time, we choose the log-normal distribution Jaynes (2003), which has a continuous positive support. Under the biological clock, any two time increments, such as $t_\Delta - t_0$ and $t_{s+\Delta} - t_s$, are jointly lognormal. When the time increments are positively correlated, they will also be positive quadrant dependent (this follows from Pitt (1982)), and hence, the corresponding discount factors will be positively correlated as well.

To make the example concrete, consider the choice between a payoff of 1 at time 0, and $1 + K$ at a delay (in calendar time) of $\Delta = 0.5$. We compare this to a pair of later options equally distanced in calendar time, at $s = 2$ and at $s + \Delta = 2.5$. In biological time, these events are at $t_0 = 0$ and t_Δ for the first pair, t_s and $t_{s+\Delta}$, for the second pair. The time of payoff delivery under the biological clock is a random variable.

We obtain the values of the options by Monte Carlo sampling. To generate the samples, we consider n *increments* in biological time, $\tau_1, \cdots, \tau_n$, that correspond to time *increments* of 0.5 units in calendar time. The n increments in biological time are drawn from a multivariate lognormal distribution with common mean 0.5 and unit variance. The correlations between the increments are positive, but decrease exponentially as they are farther apart in time. We encode the correlation structure as a covariance matrix with diagonals equal to 1, and covariances equal to $\rho, \rho^2, \cdots, \rho^{n-1}$ in the off-diagonal spots (see Appendix). We obtain instances in biological time by adding increments: $t_\Delta = \tau_1$, $t_s = \sum_{i=1}^{4} \tau_i$, and $t_{s+\Delta} = \sum_{i=1}^{5} \tau_i$. These formulae reflect the fact that the biological expiration time of the delayed early option occurs after one increment, while the two later options mature after 4 and 5 increments, respectively.

For the first option pair, the value of the immediate option is 1, and the option with payoff at (calendar

time) $\Delta = 0.5$, is valued at $E[e^{-t_\Delta}(1+K)]$, assuming a unit discount rate (in biological time). Monte Carlo analysis based on $N(=10^6)$ samples of t_Δ estimates this to be $\frac{1}{N}\sum_{i=1}^{N} e^{-t^i_\Delta}(1+K)$. For $K = 0.78$, we find that $E[e^{-t_\Delta}(1+K)] \approx 1$, i.e. the decision-maker is approximately indifferent between immediate delivery of \$1 and a payoff of \$1.78 after a delay of $\Delta = 0.5$ units of calendar time.

For the pair of options at the more distant future, the values of the early and later options are estimated to remain approximately equal when time increments are independent, i.e., $\rho = 0$ ($E[e^{-t_s}(1)] \approx E[e^{-t_{s+\Delta}}(1+K)] \approx 0.056$). When time increments are positively correlated, i.e., $\rho > 0$, the later option is valued more highly, in accordance with our Theorem. For example, when $\rho = 0.5$, the early option has value $E[e^{-t_s}(1)] \approx 0.091$, and the later option has value $E[e^{-t_{s+\Delta}}(1+K)] \approx 0.104$. So the decision-maker prefers to wait to receive \$1.78 later, while he was indifferent between an immediate \$1 and \$1.78 after an equally long delay of $\Delta = 0.5$. We thus have obtained a temporal inconsistency.

We can obtain a discounting curve by evaluating payoffs at various delays, as in the previous example. We generated $n(=10)$ increments, $\tau_1, \cdots, \tau_n$, of length $\Delta = 0.5$. The time indicated by the biological clock at calendar time s is given by $t_s = \sum_{i=1}^{s} \tau_i$. The value of a payoff of \$1 at time 0 is 1, and at (calendar) time s is $E[e^{-t_s}(1)]$ (we continue to use a unit discount rate.) The values obtained for each point in calendar time can then be compared to valuation with hyperbolic discounting (in calendar time), assuming a discount factor is $1/(1+ks)$. We find the best-fitting value of k by minimizing the squared error between the theoretical values under the hyperbolic function, and that generated by our Monte Carlo procedure. Similarly, we can also obtain a comparison with valuation (in calendar time) assuming exponential discounting, where the discount factor equals $e^{-\delta s}$. The discount rate δ is also obtained by minimizing the squared error.

We are interested, in particular, in the effect of the correlation parameter ρ of the biological clock on the shape of the discounting curve. When autocorrelation equals zero ($\rho = 0$), the discounting curve is pretty much exponential, as shown in Figure 3.7.1 (Top). The best-fitting exponential curve has $\delta = 0.3$; this differs from the true discount rate (1) because the latter only applies to biological time. The hyperbolic curve has $k = 2.75$, but its fit is worse. Figure 3.7.1 (Bottom) illustrates how, when the autocorrelations are very high ($\rho = 0.97$), hyperbolic discounting provides the better fit. The best-fitting exponential discount rate equals $\delta = 0.45$, and the hyperbolic discount rate is estimated at $k = 1.75$.

Variability in the mapping from calendar to biological time also plays a role. In the limit, when the biological clock is accurate (i.e., the mapping is deterministic, and, because we assume a constant speed for the biological clock, linear), we of course obtain exponential discounting in calendar time. Appendix 3 shows how variability induces increased convexity in the discounting curve.

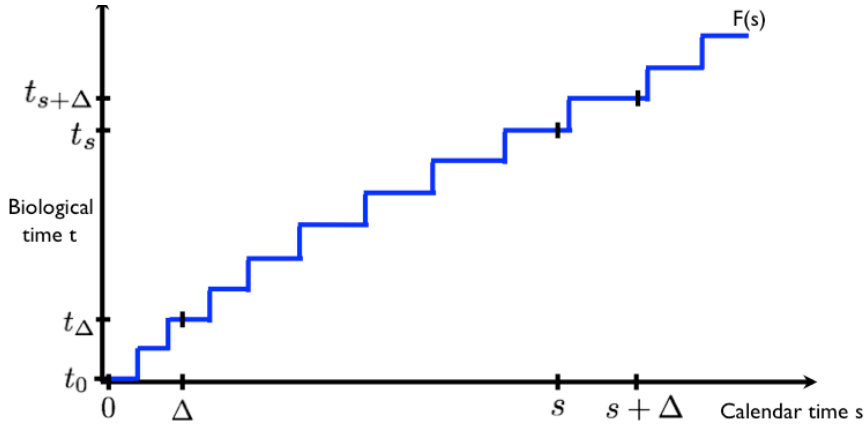


Figure 3.6: The calendar time and biological time evolve at different rates. Two equal intervals $(0, \Delta)$ and $(s, s+\Delta)$ in calendar time (horizontal axis) translate into unequal intervals (t_0, t_Δ) and $(t_s, t_{s+\Delta})$ in biological time (vertical axis). The function $F(\cdot)$ depicts one possible realization of the (stochastic) transformation from calendar to biological time.

3.7.2 Discussion

Time inconsistencies, like the ones that led to modeling time preferences with hyperbolic discounting, arise when the biological clock advances randomly in calendar time, and increments in biological time are positively dependent. When measured in calendar time, discounting becomes increasingly hyperbolic as the biological clock becomes more highly correlated and more variable.

Prior to our result, hyperbolic discounting emerged in a normative (i.e., fully rational) model because discount rates were assumed to be stochastic Farmer and Geneakopolos (2009). Our rational explanation of hyperbolic discounting does not rely on random discounting, but on randomness in the transformation between calendar time (which determines payoff times) and biological time (which is relevant for decision making). The two explanations can be shown to be related mathematically, but they are biologically very different. Specifically, stochastic time perception is biologically plausible, while stochastic discounting is rarely considered outside the arcane world of mathematical finance. An exception is Skog (1997).

Other rational explanations of hyperbolic discounting focus on specific forms of uncertainty about the ability of the payer to deliver the future payment (because he is bankrupted) or of the payee to receive it (because she may have deceased beforehand). When the hazard rate is stochastic, the apparent discount rate can be shown to become stochastic; see Sozou (1998); Azfar (1999); Sozou and Seymour (2003). However, payment uncertainty cannot provide a comprehensive explanation. In particular, it fails to explain hyperbolic discounting in experiments where design precludes bankruptcy and where the time horizon is too short for significant effects from sudden inability of the payee to take delivery; e.g., Kable and Glimcher (2007).

Our explanation of hyperbolic discounting assumes that discounting is exponential in biological time.

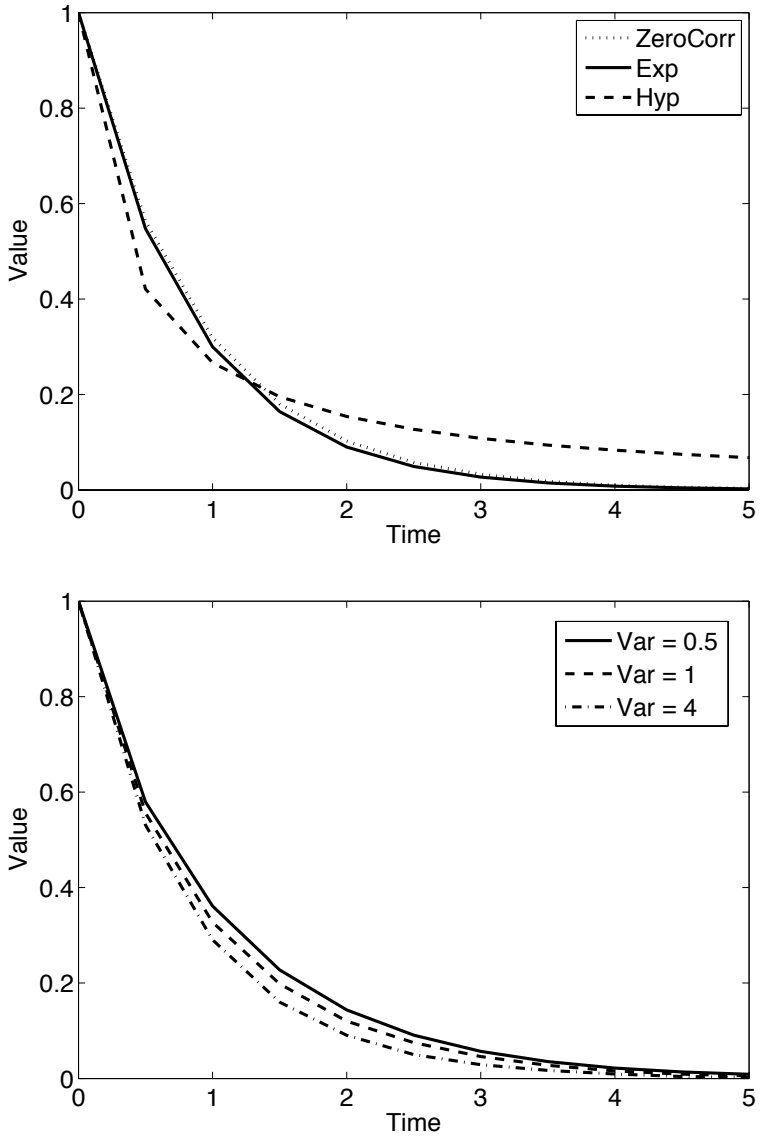


Figure 3.7: (Top) No autocorrelation of biological time ($\rho = 0$). The discounting curve in biological time is exponential with discount rate equal to 1. It generates the dotted discounting curve in calendar time ("ZeroCorr"). The best exponential ("Exp") fit in calendar time produces a discount rate of 0.30, and the best hyperbolic fit ("Hyp") has a discount rate equal to 2.75. (Bottom) Very high autocorrelation of biological time ($\rho = 0.97$). The discounting curve in biological time is exponential with discount rate equal to 1. It generates the dotted discounting curve in calendar time ("HighCorr"). The best exponential ("Exp") fit in calendar time produces a discount rate of 0.45, and the best hyperbolic fit ("Hyp") in calendar time has discount rate equal to 1.75.

Consistent with this, temporal discounting has empirically been shown to have an exponential form when subjective estimates of time elapsed are taken into account Zauberman *et al.* (2009). Other work has shown that discounting is hyperbolic if subjects perceive realizations of future events to be uncertain Dasgupta and Maskin (2005).

The importance of perceived time in discounting has been pointed out before Kim and Zauberman (2009); Nakahara and Kaveri (2010), but because random time changes were never considered, some type of misperception had to be invoked to generate hyperbolic discounting and the associated choice inconsistencies. Specifically, in Kim and Zauberman (2009); Nakahara and Kaveri (2010), the mapping from calendar to biological time is concave, so that increments farther in calendar time are expected (under the biological clock) to become shorter, inconsistent with the actual experience once the future arrives. In contrast, in our case, the speed of the biological clock is in tune with the calendar time, on average. Our approach relies on variability in the estimates.

Still, we can emulate the misperception of Kim and Zauberman (2009); Nakahara and Kaveri (2010) by increasing the expected speed of the biological clock for time farther into the future, or equivalently, decreasing the drift in the mapping from calendar to biological time. This generates concavity in the (random) mapping from calendar to biological time, and hyperbolic discounting adequately captures the resulting intertemporal choices; see Appendix 4. Positive dependence in the biological clock is no longer needed; nor is variability. As such, our framework encompasses explanations that rely on concavity in the mapping from calendar to biological time.

Stochasticity in time perception has long been accepted in psychology. Gibbon *et al.* (1984), e.g., uses a random clock process to explain response accuracy in animal timing tasks. Consideration of temporal dependence of the biological clock is novel, however, and may elucidate timing anomalies that an independent clock cannot explain Machado and Keen (1999). Positive temporal correlation in the internal clock is neurobiologically plausible; it may be supported by the positive autocorrelations recently discovered in human brain activity oscillations, displaying slow decreases even over thousands of cycles Linkenkaer-Hansen *et al.* (2001, 2004), not unlike those generated by fractional Brownian motions Mandelbrot and Van Ness (1968). It unknown to what extent this generalizes to longer time horizons, however.

Our theoretical framework provides a potentially unifying account for recorded time preferences. This is fortunate, because hyperbolic discounting is known to not be universal, with the shape (and level) of discounting changing with context Scholten and Read (2010). Context-dependence is consistent with our theory, which implies exponential discounting when the speed of the internal clock is expected to be constant, and the relation between calendar time and biological time accurate, or increments in biological time uncorrelated. Hyperbolic

discounting emerges when the biological clock exhibits temporal dependence, or when its speed is expected to decrease in the more distant future. Future research should clarify which features of the biological clock can account for the observed context-dependence of discounting. In Scholten and Read (2010), intertemporal preferences were observed to be different depending on whether a time interval is divided up, or time is extended by adding intervals. In principle, our theory could accommodate such differences, but it may require the biological clock to *not* be self-similar Mandelbrot and Van Ness (1968); that is, its temporal properties may have to change as one moves from coarser to finer sub-divisions of time.

Our theorem provides a new, unifying framework to study time perception and how it relates to impulsivity in temporal decision-making Wittmann and Paulus (2008). Our linking the phenomena of biological time and intertemporal discounting should lead to novel studies of the symptoms and causes of many disorders involving anomalous time perception, such as attention-deficit hyperactivity syndrome, borderline personality disorder, anxiety disorder, and schizophrenia.

A1. Proof of the main theorem

Theorem: Positive correlation in biological time implies temporal choice inconsistency.

Proof. To prove this theorem, we start with setting the exponential discount rate (in biological time) equal to 1. This is without loss of generality; any other discount rate would work. Now pick an amount K so that the decision maker is indifferent between an immediate (at time 0) payoff of 1 and a payoff of $1 + K$ after a delay $\Delta(> 0)$. Again without loss of generality, we set the initial biological time $t_0 = 0$ (although at times we will keep t_0 explicit, for clarity). Let t_Δ be the time that will have passed according to the biological clock by the time the calendar clock indicates Δ. Because of the preference of our decision maker, the valuations corresponding to the immediate option (left-hand side) and to the delayed option (right-hand side) are equal:

$$1 = E[e^{-(t_\Delta - t_0)}(1 + K)] \tag{3.7.1}$$

Now consider the valuation of 1 at some later time s, as well as that of $1 + K$ at the same time s plus the delay Δ. The corresponding times according to the biological clock are t_s and $t_{s+\Delta}$ respectively. The increment from s to $s + \Delta$ equals $t_{s+\Delta} - t_s$ in biological time. The (time-0) value of the payoff of 1 at (calendar time) s equals $E[e^{-t_s}]$, and the value of the payoff of $1 + K$ at $s + \Delta$ equals

$$E[e^{-t_{s+\Delta}}(1 + K)] = E[e^{-t_s}e^{-(t_{s+\Delta} - t_s)}(1 + K)] \tag{3.7.2}$$

We assume that calendar time increments are perceived to be positive quadrant dependent. Hence, $cov(e^{-t_s}, e^{-(t_{s+\Delta}-t_s)}) > 0$, or, applying the definition of covariance,

$$cov(e^{-t_s}, e^{-(t_{s+\Delta}-t_s)}) = E[e^{-t_s}e^{-(t_{s+\Delta}-t_s)}] - E[e^{-t_s}]E[e^{-(t_{s+\Delta}-t_s)}] > 0. \tag{3.7.3}$$

We can use the latter inequality to obtain a lower bound for the value of the later option:

$$E[e^{-t_s}e^{-(t_{s+\Delta}-t_s)}(1+K)] > E[e^{-t_s}]E[e^{-(t_{s+\Delta}-t_s)}(1+K)]. \tag{3.7.4}$$

Our assumption that time increments are stationary implies, in particular, that

$$E[e^{-(t_{s+\Delta}-t_s)}(1+K)] = E[e^{-(t_\Delta - t_0)}(1+K)]. \tag{3.7.5}$$

But we picked K so that the latter equals 1. Combining this with the above, we conclude that the later option is worth more than:

$$E[e^{-t_s}e^{-(t_{s+\Delta}-t_s)}(1+K)] > E[e^{-t_s}]. \tag{3.7.6}$$

But the right-hand-side is the value of the earlier option. As a result, the decision maker is no longer indifferent between the earlier and later options as she was when the earlier option was immediate; she now strictly prefers the later option, which is a time inconsistency. □

A2. Generating time increments

Biological time increments $\tau_1, \tau_2, \cdots, \tau_n$ are generated according to a multivariate lognormal distribution. The mean of the time intervals is fixed at 0.5. The covariance matrix encodes first-order auto-correlation, with correlation parameter ρ.

$$\begin{bmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_n \end{bmatrix} = logN\left(\begin{bmatrix} 0.5 \\ 0.5 \\ \vdots \\ 0.5 \end{bmatrix}, \begin{bmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \cdots & \rho^{n-2} \\ \vdots & & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1 \end{bmatrix} \right) \tag{3.7.7}$$

The biological time after s increments is given by: $t_s = \sum_{i=1}^{i=s} \tau_i$.

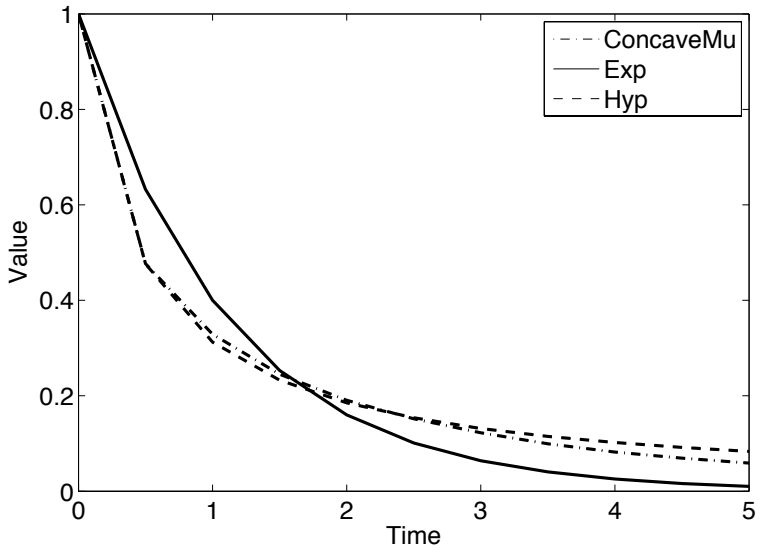


Figure 3.8: Impact of increased variability onto shape of discounting under mild autocorrelation of the biological clock ($\rho = 0.3$). As variability ("Var") increased from 0.5 over 1 to 4, the discounting function becomes more convex.

A3. Increased randomness in the mapping from calendar to biological time

Simulations are performed as for Figure 3.7.1. We set the autocorrelation of the biological clock (ρ) equal to 0.3, and increase variability (variances of the multivariate lognormal distribution of biological time intervals) from 0.50 to 1 and 4 (the middle case is the value used to generate Figure 3.7.1). Variability generates increased convexity in the discounting curve in calendar time. See Figure 3.7.2.

A4. Concave mapping from calendar time to biological time

The time increments $\tau_1, \tau_2, \cdots, \tau_n$ are generated according to a multivariate lognormal distribution. To generate concavity (on average) in the (stochastic) mapping from calendar to biological time, we let the drift in the mapping from calendar to biological time decrease with time. The expected length of the kth interval in biological time equals: $E[\tau_k] = (0.5k)^\gamma - (0.5(k-1))^\gamma$, where $\gamma < 1$ (e.g., $\gamma = 0.5$). The covariance matrix is identity, thus assuming no correlation across increments. As before, the biological time after s increments is given by: $t_s = \sum_{i=1}^{i=s} \tau_i$.

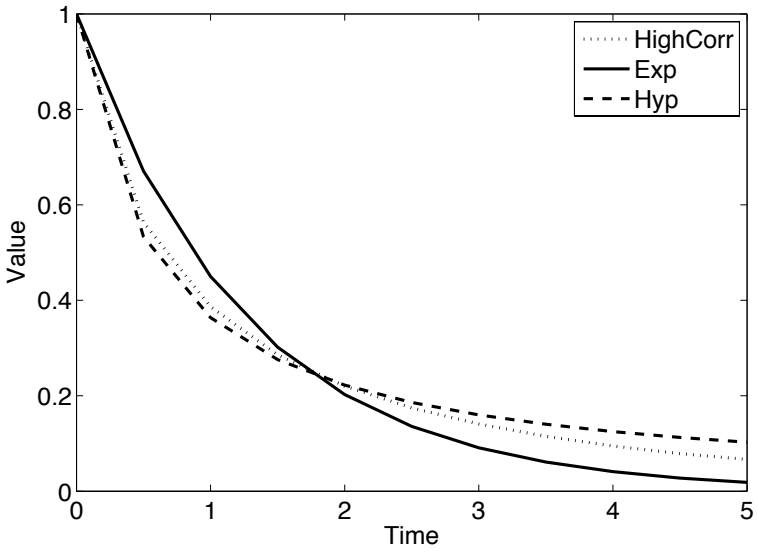


Figure 3.9: Time intervals generated from concavity in the mapping from calendar to biological time, induced by decreasing the speed of the biological clock for time intervals further into the future ($\gamma = 0.5$). The biological clock does not exhibit autocorrelation ($\rho = 0$). The discounting curve in biological time is exponential with discount rate equal to 1. The dashed-dotted line depicts the resulting discounting curve in calendar time ("Concav"). The best exponential ("Exp") fit in calendar time produces a discount rate of 0.40, and the best hyperbolic fit ("Hyp") has a discount rate equal to 2.20.

$$\begin{bmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_n \end{bmatrix} = logN \left(\begin{bmatrix} 0.5^{\gamma} \\ 1^{\gamma} - 0.5^{\gamma} \\ \vdots \\ (0.5n)^{\gamma} - (0.5(n-1))^{\gamma} \end{bmatrix}, \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \right) \tag{3.7.8}$$

We set $\gamma = 0.5$. In figure 3.7.2 we find that the discounting curve generated by the concavity in the relation between calendar and biological time is best modeled as hyperbolic.

Chapter 4

Reference Dependence and Substitution

Kahneman and Tversky (1979) introduced *Prospect Theory* to model the phenomenon, which was tested in controlled laboratory experiments, where subjects were willing to give up potential gains, to avoid suffering potential, comparatively smaller losses. This "loss-aversion" depended on a reference level (typically 0 dollars in these experiments).

Reference dependence was also discussed in the context of consumer behaviour by Thaler (1980), whereby consumers form expectations in the price or quality of commonly encountered goods, and new goods are evaluated against this reference level. Consumers are also "loss-averse" in the sense that they tend to react more strongly towards increases in price from their reference point (a "loss") compared to relative gains (Thaler (1980)). This suggests a plausible explanation for why goods are always advertised at a discounted price, and never at a premium.

Koszegi and Rabin (2006) advanced a theoretical model for reference dependence, where a person's reference point is her rational expectations, determined in a personal equilibrium consistent with her optimal behaviour given expectations. Recently, Rosato (2013) developed a model of how sellers can manipulate consumer's reference prices for substitute items. It looked at the phenomenon from the firm's point of view, and did not offer a quantitative model to measure consumer loss aversion effects among substitute products.

Studies have also noted reference dependence and loss aversion in field data. Using housing market data, Genesove and Mayer (2001) found that home sellers are averse to losses relative to their reference price - the purchase price of the home, even if the expected selling prices are going down. Camerer *et al.* (1997) and Crawford and Meng (2011) also noted this effect in a labor market for taxicab drivers, who tended to stop working when their earnings reached a reference level, even with higher demand for labor.

In marketing science, there is substantial literature on reference prices, with early references to the concept by Monroe (1973). Modeling and behavioural research has focused on the *formation* (Alba *et al.* (1999)), *recall* through internal and external processes (Mazumdar and Papatla (2000)), and the *effects* of reference

pricing (Greenleaf (1995)).

Reference price effects have been established in several empirical studies (Meyer and Johnson (1995), Kalyanaram and Winer (1995)). One common approach of modeling reference price effects (Briesch *et al.* (1997)) treats price changes as "sticker shock", with the reference price expectations formed from memory, based on the prices from the past few periods.

Estimates from scanner panel data show substantial consumer loss aversion, even after accounting for reference price heterogeneity (Bell and Lattin (2000)). Loss averse consumers make substantially more purchases when prices are discounted against their reference price, and make disproportionately less purchases when prices are raised, than can be accounted for by price elasticity. However, this does not uniquely identify loss aversion as the cause of this behaviour. An alternate explanation is that opportunistic buyers who can anticipate their future demands, "stockpile" items that are on temporary sale (see Bell *et al.* (1999)). Demand can dry up after discounts are lifted since consumers have already stocked up on their desired goods.

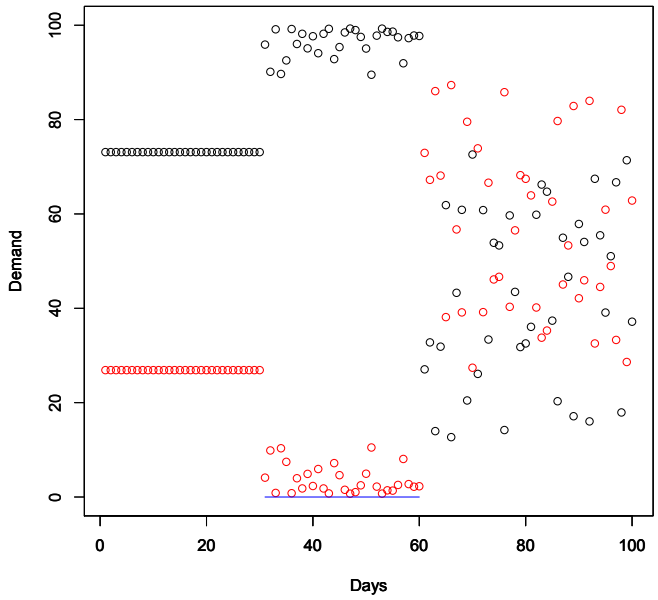


Figure 4.1: Simulated effect of loss aversion and substitution with each point a different sample of loss aversion parameter λ (sampled uniformly between 1 and 4). When higher selling item (black dots), goes on sale (periods indicated by blue line), quantity increases. When sale ends, the quantities sold decrease whilst, due to loss aversion, the relative utility of a lower selling substitute (red dots) increases. At each period the loss aversion parameter is sampled from a distribution.

We test a prediction made by the reference dependence model that cannot be attributed to stockpiling. The relative utility of an item, once it is no longer available at a discounted rate, will go down in comparison to its substitute which did not experience a similar end of sale, as illustrated in figure 4.1. Rational consumers who stockpile, buy more of an item when it is discounted, and *both* the item and its substitute will have lower sales

after the discount period is over. In section 4.3, we develop a discrete choice structural model that incorporates reference dependence to estimate consumer loss aversion from our retail dataset, and the substitution effect provides the desired identification.

Emerging evidence shows that behavioural effects are more pronounced among inexperienced players, and behaviour tends to become more rational through experience. In laboratory settings, Camerer (2003) notes that subjects become more rational through learning and behaviour starts converging towards Nash equilibrium in games, such as p-Beauty and prisoner's dilemma. In a field setting, Gneezy and List (2006) found that estimates of endowment effect and loss aversion decreased among experienced card traders in a card-trading league, and also among laboratory subjects after they gained experience. Camerer *et al.* (1997) also observed reference dependence effects in a labour market of taxi drivers, though more experienced drivers showed less of an effect. Genesove and Mayer (2001) find that home owners are averse to selling their homes at a loss with respect to a reference point (their purchase price), and the effect was twice as large among owner-occupants as compared to their more experienced investor counterparts.

In section 4.2.1, we use the transaction history of individual customers in our retail dataset to contrast the effect of experience. We test whether the effects of reference dependence is different based on sales frequency (as a proxy for experience), and estimate loss aversion in both groups.

In the next section, we provide details of our retail dataset, followed by motivating analysis for reference and state dependence. We specify our model in section 4.3, and estimate and identify effect parameters from the retail dataset. We conclude the paper with a discussion of implications for competitive pricing strategy.

4.1 Dataset

We obtained detailed consumer transaction records for the online sales division of a large retailer of hardware and household products. An example of products available is shown in Figure 4.2. The dataset contained 587 days of records from April 29th, 2010 to December 10th, 2011. 6039 items were available, with a total sales of over 100 million US Dollars. Over 1.2 million orders, or individual transactions, were recorded for 3.3 million items sold. Among the items, with individually assigned stock keeping units (SKUs), the distribution of sales is highly skewed, as shown in figure 4.3, with the top 1000 SKUs representing the majority of sales. We focus our analysis on the top 1000 ranked SKUs (according to sales), as data for the less popular items are sparse.

The items were frequently discounted. The total discounts per day, plotted in table 4.4, show that the volume of discounted sales were significantly higher in the last 8 months (days 348 to 587). The fraction of the days that an item was on sale, in order of sales rank, is shown in table 4.4. Many of the items were on sale for more than 50% of the days. For our analysis, to test the effect of a sale, we restrict the analysis to those

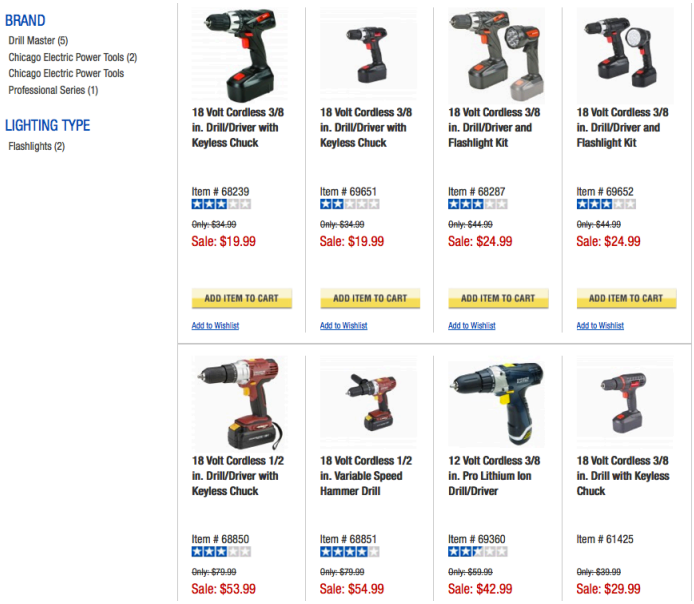


Figure 4.2: A snapshot of the retailer's website powerdrills. Consumers can easily look at the price of the item of interest, whether it is currently at a discount, as well as the price and discount status of substitutes.

items that were on sale less than 50 per cent of the time.

We extracted unique customer identification information from the transaction records. There were 861,000 unique customers, including large establishments. The total amount of purchases by each customer shows power law characteristics, with the majority of items being purchased by relatively small number of consumers, as shown on a log scale in figure 4.3. For our analysis of experience effects in section 4.2.1, we separate customers into "experienced" and "inexperienced" buyers based on total purchases.

4.2 Motivating Analysis

If consumers are loss-averse, with respect to a reference price, then discounts will have disproportionate impact on demand, beyond response to a price decrease. Furthermore, once the discount period is over, there should be a disproportionately large decrease in demand, beyond that explained by the price increase. For our initial analysis, we use a regression model that includes a dummy variable for days that an item was discounted. We also included dummy variables that indicated if the last sale ended 1, 2, or n days ago. The regression model for the quantity of item i sold at time t is given by:

$$Q_{it} = \beta_0 + \beta_{price} P_{it} + \beta_{month} Month_t + \beta_{sale} Sale_{it} + \beta_{saleend(n)} SaleEnd(t-n)_{it} \quad (4.2.1)$$

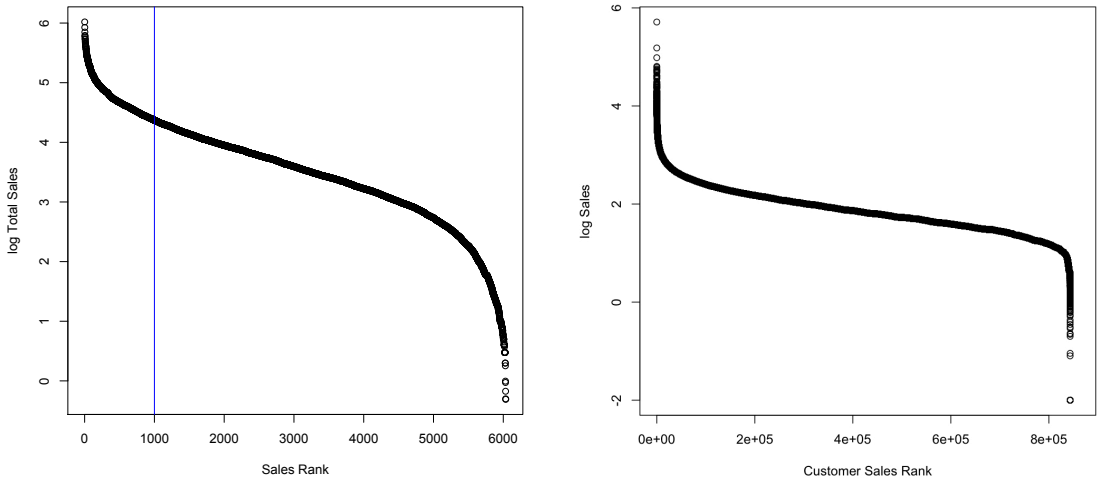


Figure 4.3: The total sales, in log scale, of all 6039 SKUs ranked (left). The total sales made by customers in log scale, in ranked order (right).

where the variable P_{it} is the price of item i at time t, $Month_t$ is an indicator of the month at time t, $OnSale_{it} = 1$ if item is on sale at time t and 0 otherwise, and $SaleEnded(t-n)$ are n dummy variables indicating whether the last sale ended n days ago (from time t).

We run the regression using the last 8 months of transactions on the top 1000 items that are discounted less than 50% of the days. Figure 4.5 shows the mean estimate (with standard errors) of the dummy variables, for sale and end of sales, for the top 1000 items.

This regression analysis reveals that being available at a discount has a significantly positive effect on sales, beyond the reduction in price. The 1-8 days following a sale shows a significant dip in sales, with the effect disappearing after the 12th day.

A likely explanation for this behaviour is consumers' loss aversion with respect to their reference price expectations. When an item is on discount, a loss-averse consumer would advance their purchase decision in order to not experience the "loss" after the temporary discounting period is over. Similarly, a loss-averse consumer would experience a "loss" when the item is no longer available at a discount, which prevents him from buying the item. The duration of this effect would last until the consumer forms new expectations for reference price, as proposed by memory-based reference price formation models (Briesch *et al.* (1997)).

A criticism of the reference-price explanation of this phenomenon is that rational consumers, who are aware that items are sometimes available at discounted prices, make opportunistic purchases during these discount windows. The drop in sales after a discount is lifted can be attributed to this "stockpiling effect" (Bell *et al.* (1999)).

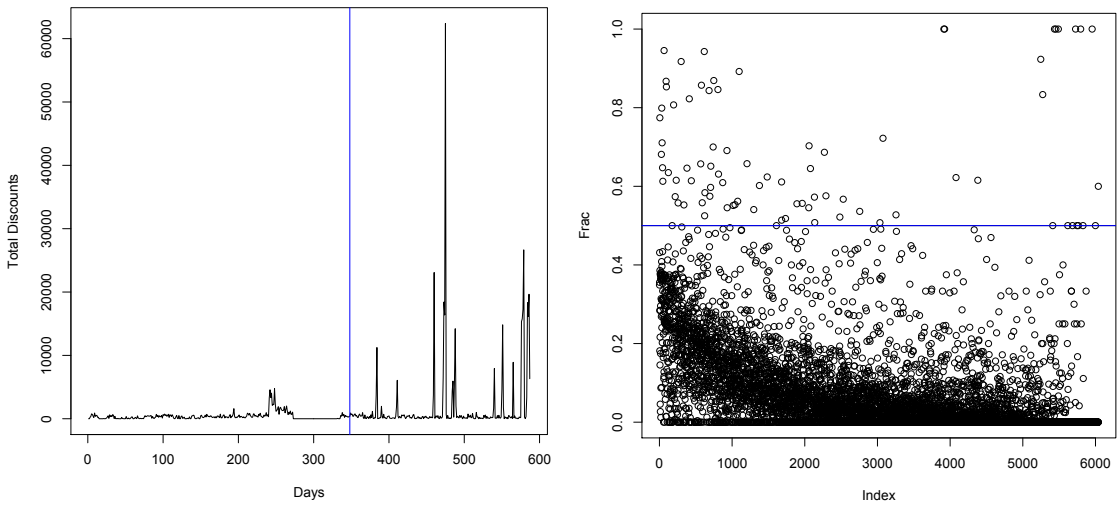


Figure 4.4: The total amount of discounts offered per day over 19.5 months (left). The fraction of days the items (ordered according to sales rank) are on sale in the last 8 months (right).

A testable prediction made by the reference dependence model, is that the relative utility of an item, once it is no longer available at a discounted rate, will go down in comparison to its substitute which did not experience a similar end of sale, as illustrated in figure 4.1. This prediction goes against the stockpiling explanation, which predicts that consumers will buy more of an item when it is discounted, and *both* the item and its substitute will have lower sales after the discount period is over[1].

Information on substitute products are available for a subset of the SKUs. While it is feasible to define a substitute as the SKU with high negative price cross-elasticity, this tends to exaggerate the substitution effect and resulting loss aversion, and also produce erroneous matches due to the large number of SKUs and daily variability in sales. We instead obtained a list of products and their substitutes from the e-Commerce source. This list is typically used by the retailers to classify similar items for a web search on the retailer's website[2] (sample shown in figure 4.2).

We show an illustrative example in figure 4.6, of the sale of two substitutable products: $50 and $100 gift cards. As expected, the sale for $100 gift card increases when a discount is offered (dots at the bottom indicate discount period) and go down once the discount is over. More interestingly, once the discount period ends, the

[1]Consider two substitutes, red and blue hammers, that are identical in all other properties except for their colour, and both sell the same number of units a day for the same price. If the red hammer is sold at a discount of 10% its sales go up during the sale, and the sales for blue hammers go down. Once the red hammer is back at its regular price, the reference dependence model says that its utility will go down relative to the blue hammer, and the sales of the substitute will be relatively higher. The "stockpiling" explanation, however, suggests that consumers will stock up on red hammers during the sale period, so after the sale, sales of red and blue hammers will be lower, and there should be no relative difference in sales.

[2]To identify a product's substitute, we perform a string-match using the product name and then remove matches that are obviously wrong, e.g. if "drill" returns "cordless electric drill" and "fire drill dvd". An alternative method is to select items by searching the retailer's website for the item and category. For more robustness, we have requested the retailer to supply us a list of substitutes that is used for placement and stocking purposes.

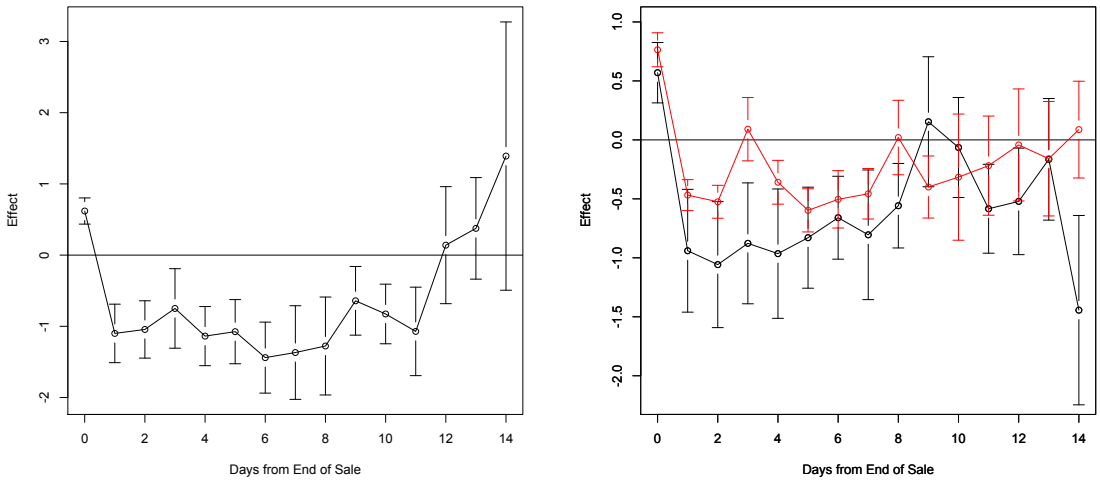


Figure 4.5: (left) The effect of Discounts (day 0) and days since the last discount period for the top 1000 SKUs. (right) The effect for "experienced" customers (red line) and "inexperienced" customers (black line).

sale for $50 gift cards go up. The effects, captured using dummy variables[3] shows a previously noted pattern for the $100 gift card: sales go up with discounts (day 0) and drop afterwards. For the substitute $50 gift card, sales increase 1 to 4 days after the discount for $100 card is lifted (and no concurrent discounts are offered for the $50 card).

4.2.1 Behaviour of Experienced Consumers

There are about 861 thousand unique customers in our dataset, with the distribution of total purchases shown on a log scale in figure 4.3. Some of the large customers are institutions or secondary retailers fulfill their inventory online. We eliminate 143 customers whose total orders exceed ten thousand dollars, and 17800 customers whose total orders are less than one dollar.

We separate the customers into quartiles based on the total number of items bought. We run the regression with the top quartile, the "experienced" group, and the bottom quartile, the "inexperienced" group, equation 4.2.1, from section 4.2 on each group, and show the results in figure 4.5. We find that both groups show a "Sales" effect, with increased sales when discounts are offered, and a dip in sales when discounts end. However, the effect lasts longer among the "inexperienced" group who would also show higher values of λ if estimated from this effect alone.

In the following section, we will test the substitution effect, and estimate values of λ for these two groups,

[3] We use a regression model as in equation 4.2.1, $Q_{it} = \beta_0 + \beta_{price}P_{it} + \beta_{month}Month_t + \beta_{sale}Sale_{it} + \beta_{saleend(n)}SaleEnd(t-n)_{it} + \beta_{saleendsub(n)}SaleEndSub(t-n)_{it}$, with variables $SaleEndSub(t-n)_{it}$ indicating that the sale for the substitute ended n days ago.

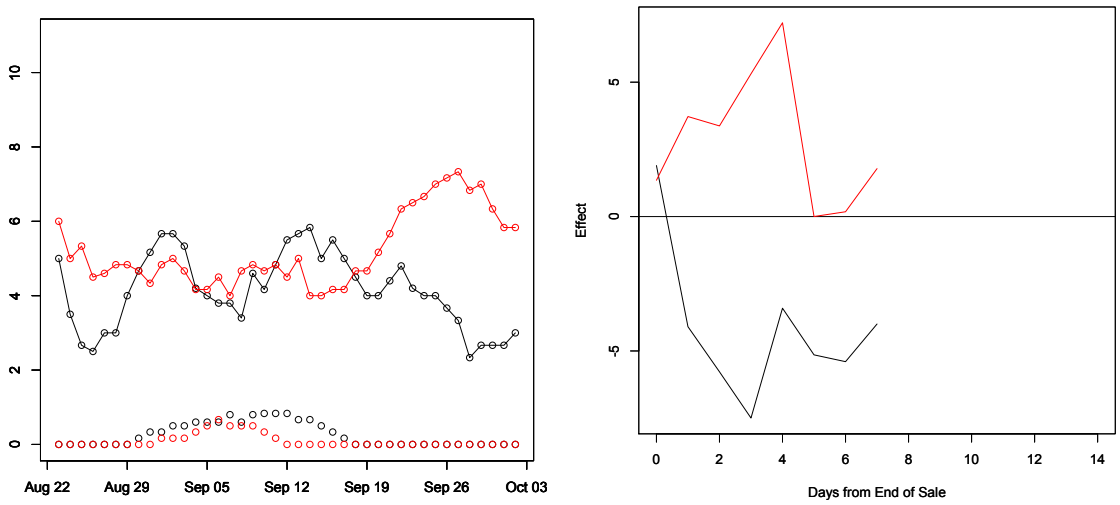


Figure 4.6: The $100 gift card has a spurt in sales during sale periods (indicated by black dots). Once the sale is removed, the quantities sold decline whilst the substitute, $50 gift card, has a greater increase in sales. On the right: The effect on sales since the day the last sale ended (black line). A sale ending for the substitute (red line) has a positive effect on the substitute. The quantities and discount periods are smoothed using a 5-day window.

using a structural model.

4.3 Structural Model

Let a product have attributes defined by a vector $\vec{a}$, and price p. The consumer assigns a value to the product based on attributes and his unobserved characteristics, represented by parameters θ. If the utility, given by the expected difference between value and price, is positive, the consumer buys the item, i.e. $U = E_\theta[v(\vec{a}, \theta) - p] > 0$

We model the utility of a loss-averse consumer whose reference price level is based on their memory of the previously posted price, p_{ref}. The consumer has unobserved characteristics θ and loss-aversion, parameterized by λ. For a posted price p, the utility of the loss-averse consumer is given by:

$$U_\lambda = E_\theta[v(\vec{a}, \theta) - \lambda(p - p_{ref}) - p_{ref}] \tag{4.3.1}$$

Thus, the consumer experiences a disutility of $\lambda(p - p_{ref})$ if item is sold above a reference price, $(p - p_{ref}) > 0$. If the item is sold below the reference price, i.e. $p - p_{ref} < 0$, then the consumer experiences a gain in utility of -$\lambda(p - p_{ref})$.

If $\lambda > 1$, then the consumer is loss-averse, and increases in price above reference price produces greater disutility than corresponding decreases in price below the reference point. If $\lambda = 1$, then the utility just reduces to $U_{\lambda=1} = E_\theta[v(\vec{a}, \theta) - p]$, i.e. the initial case.

If a rational, loss-averse consumer with $\lambda > 1$, is aware that a discount is only temporarily available, she projects that her future self treats the lower, discounted, rate as the reference price, p_{ref}. Knowing that she will experience a loss in the future, she advances her purchase decision to buy the item when it is on sale. This "anticipation effect" can affect the timing of the purchase with a lot more sales occuring near the end of the discounting period. We aim to explore this time-dependent utility in future work.

We model the choices probabilistically using a logit model, McFadden (1980). The probability of choosing item i, with utility $U(i)$ over substitutes $j \neq i$ with utility $U(j)$ is given by:

$$P(c_i = 1) = \frac{\exp(U(i))}{\sum_{j \neq i} \exp(U(j))} \tag{4.3.2}$$

where $P(c_i = 1)$ indicates the probability of the event that item i was purchased $(c_i = 1)$.

If $\lambda > 1$, the utility of item i, $U_\lambda(i)$ decreases if its sale ends in a previous period, relative to the utility of substitutes $U_\lambda(j)$ whose sale did not end (or is being sold at regular price). Thus, according to Equation 4.3.2, the observed sales of the item i would go down, and the sales of substitute items j would go up, following the end of the sale period.

4.3.1 Estimation

We estimate the loss aversion effect from choices: quantity of the product and its substitute purchased. Let y_{it} indicate the decision of consumer i at time period t. Without loss of generality, we'll model the decision to purchase product 1, and its substitute, product 2.

$$y = \begin{cases} 0 & \text{no purchase} \\ 1 & \text{product 1} \\ 2 & \text{product 2} \end{cases}$$

Let θ represent consumer preference parameters, and $P^j_{it}(\theta)$ denote the probability that consumer i chooses product $j \in \{1, 2\}$ at time t. Then,

$$P^j_{it}(\theta) = Pr(y_{it} = j) = \frac{\exp(U_{ijt}(\theta))}{1 + \sum_{j'=1}^{2} \exp(U_{ij't}(\theta))}, \text{ for } j = 1, 2 \tag{4.3.3}$$

By D_{it}^{j}, we denote the decision of consumer i at time t for product j.

$$D_{it}^{j} = \begin{cases} 1 & \text{if } y_{it} = j \\ 0 & \text{otherwise} \end{cases}$$

The likelihood, L_{it} of a dataset of observations of consumer i's transactions data over a time period t is:

$$L_{it} = \prod_{j=0}^{2} (P_{it}^{j}) D_{it}^{j} \tag{4.3.4}$$

The log-likelihood will be used to simplify computation:

$$\log L_{it} = \sum_{j=0}^{2} D_{it}^{j} \log P_{it}^{j} \tag{4.3.5}$$

Given preference parameters λ, the log-likelihood of a dataset of all consumers' ($i = \{1, \cdots, N\}$ transactions data is:

$$\log L_{\lambda} = \sum_{i} \sum_{j=0}^{2} D_{it}^{j} \log P_{it}^{j} \tag{4.3.6}$$

4.3.2 Identification

To identify the reference-dependence effect on substitution, we define a reference-dependent linear utility function derived from Equation 4.3.1:

$$\begin{aligned} U_{\lambda} &= v(a,\theta) - \lambda(p - p_{ref}) - p_{ref} \\ &= v(a,\theta) - \lambda p + (\lambda - 1) p_{ref} \\ &= v(a,\theta) - \lambda p + \gamma p_{ref}, \text{ where, } \gamma = \lambda - 1 \\ &= v - \lambda p + \gamma p_{ref}. \text{ (for simplicity)} \end{aligned}$$

The implications of the parameters are straightforward with this simple linear model:

- For $\lambda = 1, \gamma = 0$ and $U_{\lambda} = v - p$. So the utility is equal to the consumer suplus. If the value of the item, v is greater than the price, p the consumer will buy the item. Thus the consumer has no loss-aversion w.r.t the reference price.

- For higher values of λ, the utility decreases below the consumer surplus proportionately to the difference in current price from reference price, $p - p_{ref}$. And the utility increases above the consumer surplus as the current price drops below the reference price.

- For $\lambda = 0$, $U_\lambda = v - p_{ref}$. So the consumer is blind to the current prices, and only cares about the reference price.

Let's say the consumer's (private) value of an item is \$20. The previous price i.e. reference price is \$5, and the current price is \$15 (the price increased by \$10). The utility for a non loss-averse consumer is equal to the consumer surplus, i.e. \$5, and the consumer would buy. For a loss-averse consumer, with $\lambda = 3$, the utility is -\$5 and the consumer would not buy the item.

The parameter λ can be estimated from choice data:

$$(\hat{\lambda}) = \max_{\lambda} \log L_\lambda \tag{4.3.7}$$

From choice dataset (in a subset of consumers or time-period) we can estimate the parameters using current prices and reference prices, and test the following reference-dependence effects:

- Test of *strict* Reference Dependence: $\lambda > 1$.
- Test of *strict* Stockpiling: $\lambda = 1$.

If $\lambda = 0$, then the consumer only cares about the reference price p_{ref} and the utility is not affected at all by the current prices p (which could be infinite). Furthermore, for $\lambda < 0$, the consumer is "loss-seeking" and gains more utility from an item as its price goes higher than the reference price. Both those scenarios are absurd, and we only consider positive values of λ by estimating the transformed value, $\lambda = \exp(\kappa)$.

The effect of loss aversion on the sales of substitutes enters the picture in the probability of choice, $P(c_i)$. A loss-averse consumer would become more likely to buy the substitute after the primary good is taken off discounts, as we simulate and estimate in the next section.

4.4 Results

We first simulate our model for different degrees of loss aversion and look at the qualitative effect. We then estimate the parameter from retail data, as well as effect of experience.

4.4.1 Simulations

We simulate the choices from the discrete choice model with the utility function described in the previous section. Product 1 has a regular price of \$8, and a discounted price of \$7. The discount is applied from periods 31 to 60, and the regular price in periods 1 to 30, and after the discounting periods, from 61 to 90.

The substitute product, 2, has a regular price of \$9 in periods 1 to 90 (is not offered at discount). We assume, w.l.o.g. that the private value for a buyer is $v = \$10$; since the model is linear, the results would hold for any value of v.

For $\lambda = 1$, there is no loss aversion. Utility of product 1 at the regular price is equal to the consumer surplus, i.e. $U_1 = 2$. At the discounted price, the utility is $U_1 = 3$. For the substitute, product 2, the utility at all periods is $U_2 = 1$. As shown in Figure 4.7(left), the probability of chosing product 1, $P(C_1 = 1) = \exp(2)/(\exp(2) + \exp(1)) = 0.73$ is periods 1 to 30 and 61 to 90. In periods 31 to 60, with discounting, the probability increases to $P(C_1 = 1) = \exp(3)/(\exp(3) + \exp(1)) = 0.88$.

If $\lambda = 3$, there is a significant effect of loss aversion. From periods 1 to 30, $U_1 = 2$, while the subsitute product has utility $U_2 = 1$. Probabily of choosing 1, $P(C_1 = 1) = 0.73$. When the price is discounted, however, the consumer has a reference price, P_{ref}, equal to \$8, and current price $P = 7$. The utility of the loss averse consumer is $U_1 = 10 - 3 * 7 + 2 * 8 = 5$. The probability of chosing product 1, is substantially higher at $P(C_1 = 1) = \exp(5)/(\exp(5) + \exp(1)) = 0.98$. In periods 61 to 90, when the price rises back to the original price of \$8, the reference-dependent consumer uses the price of \$7 from the previous periods as reference price, so the utility decreases to $U_1 = 10 - 3 * 8 + 2 * 7 = 0$. The utility for product 1, is less than the utility for product 2, and $P(C_1 = 1) = \exp(0)/(\exp(0) + \exp(1)) = 0.27$. The loss-averse consumer is now *more likely* to buy product 2 in periods 61 to 90.

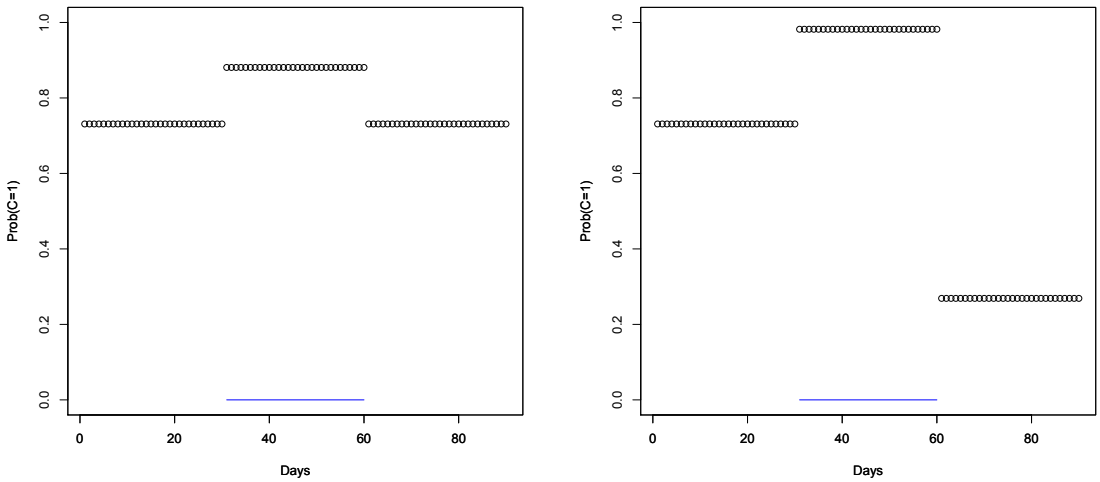


Figure 4.7: The probability of choosing first item, $P(C_i = 1)$, for no loss aversion, $\lambda = 1$ on left. $P(C_i = 1)$ for loss aversion coefficient $\lambda = 3$ on right.

We then ran the simulation where, at each period, λ is sampled from a uniform distribution between 1 (no loss aversion) and 4. For clarity, we multiply the probabilities of choosing each item by 100 units, and plot the

per period demand in Figure 4.1.

4.4.2 Estimation from Retail Data

Illustrative Example

To illustrate the method of estimating the reference dependence (loss aversion) parameter, λ, we consider the pair of substitutable products in the previous section: the \$50 and \$100 gift cards (shown in figure 4.6). Due to the daily fluctuations in sales quantity and price, we apply a "smoothing window" of one week. The reference price, p_{ref} is the average of the previous week's prices, and the current quantity, q_t is the average of q_t to $q_t + 6$. Note: The results are not different for window lengths between 3 to 10 days, and one week is a representative length of time for online sales.

In addition to the loss aversion parameter, λ, the private values for the two products, v_1 and v_2 are unknown. We estimate the likelihood of the data under different parameter values of λ, v_1, v_2. λ ranges between 0.1 and 5.0 in increments of 0.1 (since $\lambda = \exp(\kappa)$, we have κ range between $\log(0.1)$ and $\log(5.0)$. v_1 and v_2 is allowed to range between 0 and 200 (upper range is twice the maximum price), in increments of 10. We then take the maximum likelihood estimate (MLE) jointly over λ, v_1, v_2.

For the pair of \$50 and \$100 gift cards, the MLE of $v_1 = 60$ and $v_2 = 110$. The private values are higher than the median price of the two products, \$50 and \$100 respectively. This is consistent with the utility model where consumers will buy the product if their private values are higher than the price of the product. MLE of $\lambda = 1.5$, with a log-likelihood of -1108.32; in contrast the log-likelihood by fixing $\lambda = 1$ was -1212.02; with a log-Ratio $\big(\frac{\log L(\lambda^*)}{\log L(\lambda = 1)}\big)$ of 0.91. Thus we find evidence of reference dependence and the effect on substitutes on these pair of substitutable products.

Estimation on Dataset

We then applied the estimation method across all products that are within the top 1000 SKUs in terms of sales. For each pair, we jointly estimated the private values v_1, v_2 and loss aversion parameter λ. v_1 and v_2 ranged between 0 and twice the maximum price of the item, and $\lambda \in \{0.1, \cdots, 5.0\}$.

We contrasted the estimated values of λ among the substitute products, as well as the estimate among the non-substitutable products. Table 4.1 shows estimated values for the top 3 products (in terms of total sales) and their substitutes, as well as some that are not substitutes. We also show the ratio of the log-Likelihood of the pair under the multinomial choice model for the MLE λ value, and for $\lambda = 1$ (logRatio $= \frac{\log L(\lambda^*)}{\log L(\lambda = 1)}$).

The conclusion that we draw regarding the estimates of λ for substitute and non-substitute products is critically based on how we define substitutes. We have avoided debate on the definition by using an internal list used by the retailer. The retailer maintains this list to suggest products that are similar on the website, as

well as internal decisions on which products to offer for promotions. Alternate means of finding substitutes ourselves, such as calculating cross-elasticity, is prone to noise and identifies pairs as substitutes that clearly are not[4]. In the absence of a list from a list from the retailer, we could have "crowd-sourced" this identification to random shoppers on the internet using Amazon mechanical turk or Google Consumer Surveys[5].

The list from the retailer identified 3673 substitutes among the top 1000 SKUs. The median number of substitutes per SKU was 3. The mean value of estimated λ was 1.43 among the substitutes, with a standard error of 0.014 ($n = 3673$). The mean value of λ among the 495,827 non-substitutes was 1.06, with a standard error of 0.008 ($n = 495827$). Figure 4.9 shows a bar graph of the λ values for the two samples.

Note that the estimated value of λ varies considerably, among both the samples. The highest estimate of λ among the non-substitutes was 5, while the highest value of λ was also 5 among the substitutes. This variability is due to the inherent noisiness (daily variability) of eCommerce sales data. In future work, we aim to develop robust estimation methods that deal with the daily variation in the sales data. We applied the two-sample t-test to test the null hypothesis that the λ estimates of the substitute samples has the same value as the λ estimates of the non-substitute samples. We reject the null hypothesis, and accept the alternative that the mean of the former samples is greater, with a t-statistic of 7.22, and p-value of 8.110^{-12}. We conclude that the λ estimates of the substitutes is significantly higher than the non-substitute samples.

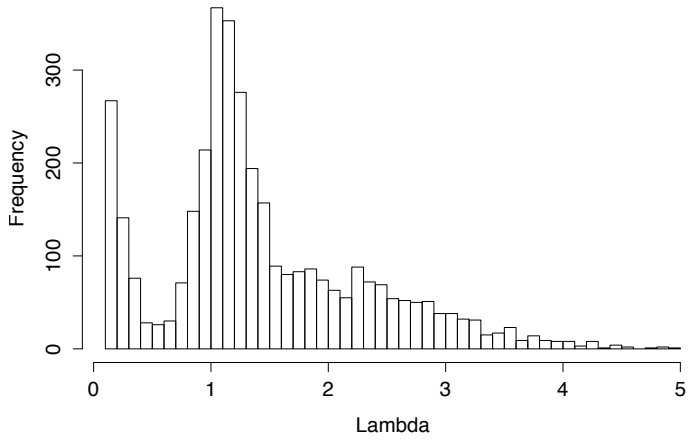


Figure 4.8: A histogram of the estimated values of λ for all 3673 substitute pairs.

Figure 4.8 shows the distribution of λ for all the 3673 substitute pairs. We find a cluster of values above $\lambda = 1$, and a smaller cluster near $\lambda = 0.1$. The average difference between the private values v_1 and v_2 and the average price for the item is $\$11.65(\pm 6.11)$. This is reasonble since under the utility model the purchase

[4]We computed the cross-elasticity of the top selling SKU, "45 Watt Solar Panel", against the top 1000 SKUs. The two SKUs that had the highest cross-elasticity score were: "6-inch Digital Calipers" and "1200 lb. capacity air motorcycle lift".

[5]Amazon Mechanical Turk: https://www.mturk.com/mturk.
Google Consumer Surveys: https://www.google.com/insights/consumersurveys.

SKU_1	SKU_2	λ	logRatio
45 Watt Solar Panel Kit	3 Piece Decorative Solar Light Set	0.9	0.99
45 Watt Solar Panel Kit	10 Piece Stainless Steel Solar Light Set	1.5	0.88
45 Watt Solar Panel Kit	80 Watt Monocrystalline Solar Panel	1.4	0.86
45 Watt Solar Panel Kit	Saw Mill with 7 HP Gas Engine	1.0	1.00
45 Watt Solar Panel Kit	Automatic Battery Float Charger	0.7	0.93
45 Watt Solar Panel Kit	4-1/2 Angle Grinder	1.0	0.93
90 Amp Flux Wire Welder	120 Amp Arc Welder	1.2	0.92
90 Amp Flux Wire Welder	80 Amp Inverter Arc Welder	1.1	0.91
90 Amp Flux Wire Welder	115 Volt Spot Welder	1.6	0.81
90 Amp Flux Wire Welder	230 Volt Spot Welder	1.8	0.71
90 Amp Flux Wire Welder	135 AC/105 DC Amp Power Arc Welder	1.5	0.93
90 Amp Flux Wire Welder	100 Amp Flux Wire Welder	1.3	0.89
90 Amp Flux Wire Welder	240 Volt Inverter Arc/TIG Welder	0.9	0.96
90 Amp Flux Wire Welder	Saw Mill with 7 HP Gas Engine	1.0	1.00
90 Amp Flux Wire Welder	Color Security Camera w/ Night Vision	1.0	1.00
90 Amp Flux Wire Welder	4-1/2 Angle Grinder	1.0	1.00
11 Drawer Roller Cabinet	13 Drawer Red Industrial Quality Roller Cabinet	2.2	0.79
11 Drawer Roller Cabinet	Red Roller Cabinet Top Chest	1.8	0.68
11 Drawer Roller Cabinet	8 Drawer Roller Cabinet with 8 Drawer Top Chest	1.6	0.66
11 Drawer Roller Cabinet	20 Drawer Roller Cabinet - Red Glossy Finish	2.4	0.62
11 Drawer Roller Cabinet	13 Drawer Black Onyx Roller Cabinet	1.3	0.90
11 Drawer Roller Cabinet	40 Lb. Capacity Floor Blast Cabinet	1.0	1.00
11 Drawer Roller Cabinet	Heavy Duty Milling/Drilling Machine	0.1	0.96
11 Drawer Roller Cabinet	Semiautomatic Tire Changer	0.1	0.96

Table 4.1: Examples of λ values for 3 top-selling SKUs along with their substitute SKUs as well as examples of non-substitute SKUs. The ratio of the log-Likelihood of the pair under the multinomial choice model for the MLE λ value, and for $\lambda = 1$ is also shown; logRatio $= \dfrac{\log L(\lambda^*)}{\log L(\lambda = 1)}$.

only occurs when the private value is higher than the price.

4.4.3 Effect of Experience

We split the sales data for the top 1000 SKUs into two groups: the data from the top quartile (top 25%) of customers, the "experienced" group; and the data from the bottom quartile of customers (in terms of sales made on the site), the "inexperienced" group. Out of 3673 substitute pairs in the original dataset, only 2710 pairs of SKUs had sufficient data in each group. For estimation of λ we considered only the SKU pairs that have at least 25 days worth of sales data from each group of customers. We calculated the maximum likelihood estimate of λ for each of the 2710 SKU pairs separately for each group using method described above.

The mean value of estimated λ among the substitutes in the *experienced* group was 1.22, with a standard error of 0.032 ($n = 2710$). The mean value of λ among the substitutes in the *inexperienced* group was 2.16, with a standard error of 0.090 ($n = 2710$). Figure 4.9 shows a bar graph of the λ values for the two groups. The λ values of substitutes, in both the *experienced* and *inexperienced* group, is also significantly higher than

the non-substitute goods, as shown in figure 4.9.

Since there is considerable variability in the λ values for different SKU pairs, we use the two-sample t-test to test the null hypothesis that the distribution of the estimated λ values between the two groups is equal. The test rejects the null hypothesis and accepts the alternative that the mean of the latter group is higher, with t-statistic of 4.10, and p-value of $p = 0.00021$. We conclude that the λ estimates are significantly higher in the *inexperienced* group.

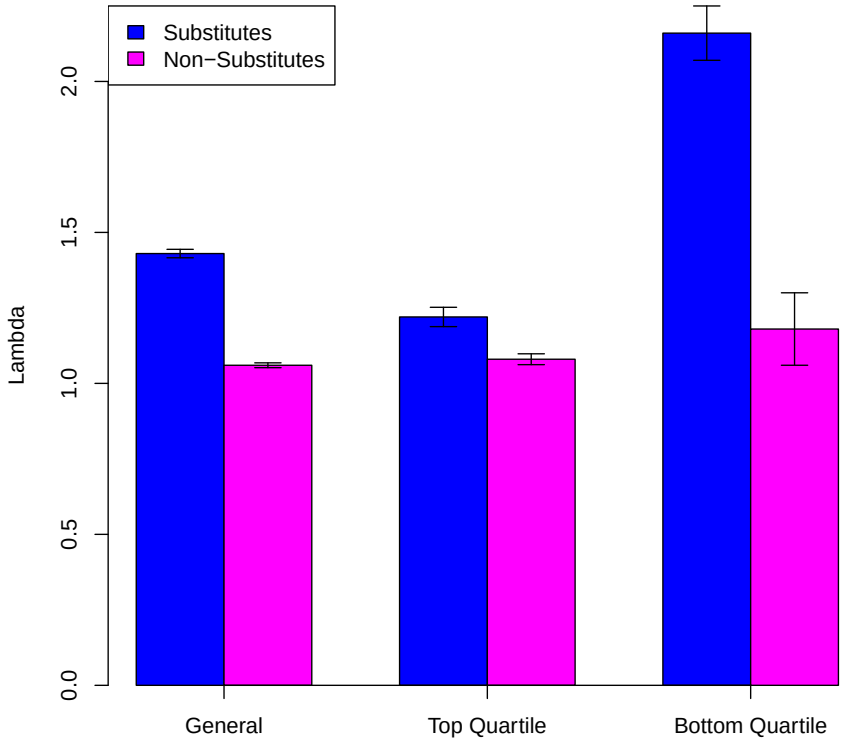


Figure 4.9: The estimates of λ (and errorbars) for substitute and non-substitute goods among the general (left), *experienced* (middle) and *inexperienced* customers (right).

The estimates of the λ for the second and third quartiles are 1.85 and 1.88 (with standard error of 0.061 and 0.087 respectively). Thus the estimates of λ for the top quartile is significantly below the estimates for the second, third and bottom quartile. An explanation for this is that the top quartile of consumers in this dataset are likely to be businesses that can purchase large quantities of the same item and stockpile.

4.5 Discussion

Increased sales of durable goods, beyond that explained by a response to price change, followed by a slump when the sales are lifted, can be attributed to reference dependence. Estimates using scanner data has shown significant loss aversion in literature. However that does not identify the effect since other factors, such as stockpiling could be occurring.

The substitution effect, whereby the relative sales of a substitute product increases when a discount is lifted, helps to identify the reference-dependence model. We found evidence of reference dependence and loss aversion in the sales of substitute goods. The average loss aversion parameter estimated on 3673 SKU pairs was significantly higher than the non-substitute SKU pairs. We also found that the effect is more pronounced among *inexperienced* consumers compared to *experienced* consumers.

A valid criticism of our findings is that in order for the consumers to exhibit loss aversion and prefer substitute products, they need to know when items are put on sale and when they are removed from sale. Since we only have data when a sale occurs, we are limited by the data in finding out if the consumer has seen the past prices before making a sale. This criticism affect reference dependent pricing literature in general (see) which relies on data to identify how customers are tracking price changes and forming reference prices. Furthermore, our estimation provides evidence against the "stockpiling" explanation which doesn't suggest this switching of preferences for substitutes.

There is sufficient evidence that in the eCommerce setting consumers track prices very frequently and do comparison shopping online. The "transaction cost" for finding prices is very low (just a mouse click away) and several online sites[6] track prices across multiple websites. The industry averages of the ratio of visitors who look at a page, to the number of purchases was 5.1% over the 2013 year (according to Internet Retailer www.internetretailer.com). The number of repeat visitors was 30-65% over the same period across the top 500 e-Retailers. This low visitor to sale ratio and high percentage of repeat visitors is universal across eCommerce sites, and should be similar for the eCommerce retailer used in our estimation as well. Furthermore, figure 4.2 shows the typical page for an item, which indicates whether an item is at a regular price or on sale, as well as similar and substitutable items and their prices along with any discounts on them. So consumers are not only aware of the price of the product (and whether it's on discount), but also the price of the substitute as well as its discount status.

Due to the daily variation in sales data, there is high variability in the estimates of λ among the different pairs of substitute SKUs. In future work, we aim to develop more robust methods of estimating loss aversion by modeling the "noise" in the sales data. Furthermore, although we consider only pairs of substitute products,

[6] Some online price comparison sites are www.pricegrabber.com, www.pronto.com

we can extend our method to estimate the reference dependence effect on a basket of goods.

The values of λ estimated in our dataset is lower than that found in laboratory experiments of reference dependence and loss aversion, which can have several explanations: (i) In lab experiments, λ is measured at an individual level among gambles that are designed to measure the effect. Instead, here we are inferring the effect from aggregate choice data. (ii) The aggregate choice data includes consumers with varying levels of experience and loss aversion. Because of the type of retailer, consumer choices can be based on industrial demands that are less sensitive to behavioural effects. We see this once we divide the dataset into customer types. Future work will look at a variety of types of items, like fashion clothing, that are more individual purchases.

Consumer reference dependence, even if observed in a subset of goods, has important implications for competitive strategy, pricing, and the timing of promotions. According to prevailing theory, firms could offer items at discounts, or have promotions, for rational consumers in order to take away sales from their competitors (Bell *et al.* (1999)). However, if consumers are loss averse, and not prone to stockpiling, then discounts and promotions have an adverse effect once the period is over, and consumers can switch to competitors' offerings.

Even if our method is not robust enough to identify substitutes from data, the estimated λ is useful in projecting the effect a price reduction or promotion can have on a product's substitutes. We have provided a quantitative model for firms to use in optimizing the timing and magnitude of price discounts on a set of SKUs. Furthermore, since the only two variables used in our model is price and sales quantity, firms can use competitor's observable sales data (even if the sales quantities are approximate) to predict the competitor's pricing on their own sales, and vice versa.

The theory presented in this paper suggests interesting experiments to run, or questions to answer if we had access to different data.

Chapter 5

Discussions

In this thesis, we presented Bayesian Rapid Optimal Adaptive Designs (BROAD) - a methological advance in adaptive experimental design. BROAD allowed us to design experiments to compare theories of risky choice in real-time experiments. Comparing behavioural theories with prior methods was difficult or infeasible due to the number of parameters which requires a very large design space. The objective criteria, EC^2 used in BROAD was proved to be adaptive submodularity which allowed us to prove theoretical guarantees of near-optimality, and write a faster implementation using accelerated greedy algorithms that led to orders to magnitude in speedups.

We also applied BROAD to also compare some key theories of time preference and found that generalized hyperbolic discounting describes subjects' behaviour the best. We also proved a theoretical result for a psychological time preference models, where positive dependence in the subjective perception of time led to hyperbolic discounting and temporal choice inconsistency.

Finally, based on the evidence for prospect theory in our lab experiments with risky choices, we sought to find evidence of prospect theory in the wild. We found evidence for loss aversion in consumers (using data from a eCommerce retailer) by testing a distinct prediction of the theory, as compared to what a standard model predicted.

At the end of the thesis there are many more open questions than answers. There are several directions that I would have naturally extended this thesis to, and several studies that could have been done. The purpose of this line of research has mainly been to broaden the scope of behavioural and experimental economics, and to push the envelope in terms of the experiments that can be done. The methods and empirical studies presented in the thesis are only the beginnings, and I briefly describe some nascent ideas that have spawned off of this research.

5.1 Future Directions

Many economic theories sacrifice realism to gain mathematical tractability. By economizing experiments, BROAD methods can be used to test complicated, realistic theories using complex designs and utilize measurements that go beyond choice data. Theories that have been vetted in the lab can also be tested and estimated with field data. This leads to faster accumulation of evidence for certain theories while eliminating those that do not have evidence from both lab and field studies.

5.1.1 New measurements in Experimental design

BROAD provides a means of including neurophysiological and psychological measurements to augment experimental design. Standard experimental economics has so far looked at choices alone. The aim of behavioural economics is to incorporate psychological realism into economic models, and knowledge of physiology should allow us to incorporate biological measurements as an added source of data. Such models would inevitably be more accurate than "as-if" models. Suppose model A and model B both fit choice data equally well, but additionally model B fits neurophysiological data. We'd claim that model B would have better realism and out-of-sample predictive power than model A. We now discuss different biological measurements, from simple response times, to complex fMRI data, that can be utilized in improving experimental designs.

Response times Response times (RTs) are simple add-on measurements for any computerized laboratory choice-based experiment. Race-to-barrier models (also called drift-diffusion) of value accumulation describe choices that vary in difficulty lead to differences in RTs (Smith and Ratcliff (2004)). For binary choice between two valued alternatives, difficult typically means that choices are close in subjective value. Response times could provide statistical information to differentiate models of valuation, if they are related to choice difficulty.

Denote the response time of observed choice a by $t(a)$. The differential subjective utility of choice a compared to unchosen alternative b under model k is $U(a|k) - U(b|k) \equiv \Delta U(a|k)$.

As a motivating example, Figure 5.1.1 illustrates a common pattern of RTs (y-axis) as a function of differential subjective value (x-axis), for one model and one subject from our risky choice experiment. The thickest inverted-U shaped line in the center represents the best quadratic fit of the actual RTs. RTs are higher for choices in which the value difference is close to zero (those are the hardest decisions and take the longest). The inverted-Us are contour plots of the same probability showing possible relations between $\Delta U(k)$ and response time t. The actual RT level in one trial is shown by a horizontal line cutting across the figure.

A theory H_a which is accurate (the actual choice has a positive utility differential according to the theory) predicts a smaller RT is likely. The intersection of the theory's $\Delta U(H_a)$ value and the observed RT lies on

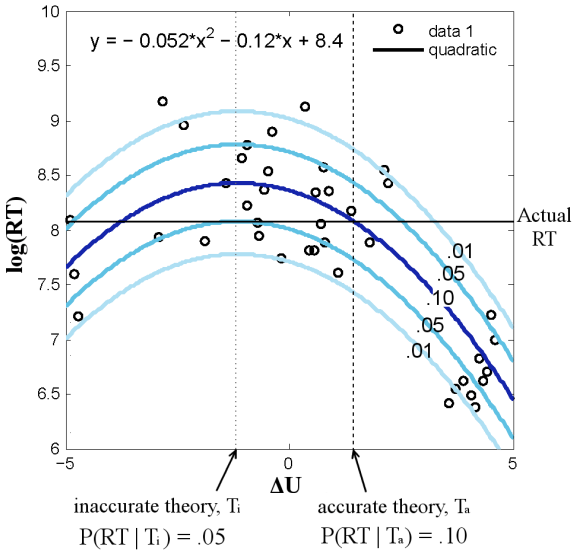


Figure 5.1: Response times (in milliseconds) in choosing between two delayed rewards in Risky Choice experiment. The x-axis is a measure of differential subjective value.

the isoquant with predicted RT probability of .10. Another theory H_i is inaccurate, since it assigns a small negative utility differential to the observed choice. Theory H_i's utility differential intersects an RT isoquant with probability of only .05. Intuitively, H_i predicts that since the $\Delta U(T_i)$ is negative and close to the RT peaks, the RT should be slow, not fast. The relative RT probabilities of .10 and .05 inferred from the theories $\Delta U(k)$ values means that the more accurate theory H_a will have higher likelihood at the expense of the inaccurate theory H_i.

Including RT information could be used in different ways. As illustrated above, it could be used after all choices are made (using terminal values of $p(k|history)$) to sharpen all the terminal estimates of model likelihood. RT information could also used for "online" updating after each trial, which requires recomputing the empirical $f(t|a)$ function after each choice-RT observation.

Intuitively, including RTs will generally "double count" a models weaknesses, and more rapidly eliminate models that fit both choices and RTs poorly. However, if there are thoughtless quick mistakes (i.e. resulting in low RT) then the method will actually compensate for them. A theory that predicts the wrong choice ($\Delta U(k) < 0$) but has a fast RT will also have a predicted RT which is fast. Therefore, the fast RT will forgive the theory, when compared to theories that predict indifference $\Delta U(t, i) \approx 0$ and hence predict a fast choice.

RTs typically fall sharply over trials (in a classic power-law learning curve), as we see in Chapter 1. The average RTs were 15, 7 and 3 seconds on trials 2, 10 and 50. Since the first few trials are the ones in which BROAD approaches add the most information, it will be important to adjust for these learning curve effects.

Skin Conductance Response Skin Conductance Response (SCR) measures are an index of a level of arousal elicited by aversive emotions such as anxiety and fear (Figner and Murphy (2010)). Kang *et al.* (2012) empirically investigates temporal anxiety in economics decisions using two variants of a "clock game" where maximizing payoffs depend on strategically timing selling decisions. Temporal anxiety would cause participants in a dynamic version of the game to sell too early and not realize potential exponentially greater gains; the behavioural data does support the fact that subjects sell earlier in the dynamic version than in the static game. Direct measurement of SCR supported this interpretation. Measuring SCRs at the time of selling shows stronger anxiety in the dynamic game in general, and subjects who show higher anxiety (as measured by SCR) also sell relatively earlier.

Similar to RTs, SCR measurements can also be incorporated into BROAD to more rapidly eliminate models that fit choices and SCR measurements poorly. A likelihood function can map choice alternatives to differences in SCRs: we conjecture that choosing between option that are similar in utility leads to higher anxiety, and SCRs.

Eye-tracking Recently, eye-tracking has been employed to assess subjects thinking processes in various decision environments: to determine how subjects detect truth-telling or deception in sender-receiver games (Wang *et al.* (2010a)). how consumers evaluate comparatively a huge number of commodities, as in a supermarket setting (Reutskaja *et al.* (2011)); and the relationship between visual attention (as measured by eye-fixations) and valuation of commodities in choice tasks (Krajbich *et al.* (2010)). Hu *et al.* (2013) use eye-movement measurements, as auxiliary measures of subjects' beliefs, and aid in identification and estimation.

Magnetic Resonance Imaging (MRI) Increasingly a number of Neuroeconomic studies have involved Function Magnetic Resonance Imaging (fMRI) data for model comparison. Typically, there are potential number models that explain a behavioural phenomenon with similar accuracy, especially in simple tasks. Prevost *et al.* (2013) found support for model-based over model-free learning in Pavlovian learning tasks using fMRI data, and Symmonds *et al.* (2010) used fMRI data from a sequential choice task to support a sequential updating strategy over a optimal predetermined strategy.

There are at least a couple of reasons why BROAD methods are particularly suitable for fMRI studies. Scanner time is an expensive resource and subjects can experience discomfort, hence tasks cannot be very long. To allow for the BOLD signal to stabilize, several seconds of lag is included which gives time to compute new questions. BROAD methods are particularly suitable for studying complex tasks in fMRI, such as decision making in financial bubbles (De Martino *et al.* (2013)), where subjects can be paired with different groups or

shown different markets based on their previous responses.

5.1.2 Applications of BROAD

In many areas of empirical behavioral economics, one or more functional forms have been developed which appear to fit choice data comparably well. Controversies about which theories are most accurate typically persist since different people use different methods, goodness-of-fit criteria, and experimental stimuli. Application of the BROAD technique could sharpen these comparisons.

Social Preferences Mathematical models of social preference have been widely used to explore sociality of different species, gene and hormone influences, and underlying neural circuitry (Fehr and Camerer (2007)). An approach pioneered in psychology and economics (Bazerman *et al.* (1995)), is to model preferences over social outcomes as a weighted sum of an agents own outcome x_i, the outcomes of $N-1$ others x_k and some comparative function of x_i and x_k. "Inequality-averse" preferences reflect a distaste for envy (when i gets less than k) and guilt (when i gets more than k). The weights $\alpha = \beta = 0$ correspond to pure self-interest. A similar form is due to Bolton and Ockenfels (2000) where is utility is maximized at an equal share $s_i = 1/N$. A different "Rawlsitarian theory" combines self-interest with desires to help the worst-off person and to do the best for society as a whole (Charness and Rabin (2002)).

Engelmann and Strobel (2004) introduced a useful three-player paradigm to compare some of these theories. A generalization of that paradigm would utilize BROAD to find sets of three-payoff vectors which discriminate among theories best. Importantly, it has been speculated (Fehr *et al.* (2006)) that there are interpersonal differences in the emphasis placed on efficiency (the weight w_t in the Rawlsitarian approach) compared to inequality (α, β in Fehr-Schmidt and equal share preference in ERC). If so, a BROAD procedure designed to customize the choices made for each subject will have an advantage in detecting those individual differences rapidly and precisely.

Behavioural Game Theory Several solution concepts have been developed to predict patterns of behaviour in noncooperative games, including best-response and quantal-response equilibrium analyses, and level-k reasoning and cognitive hierarchy (Camerer *et al.* (2004)). Finding normal or extensive form games on which theories make very different predictions is a difficult task since the space of possible games is very large.

Different theories of learning in games have been developed, such as reinforcement, fictitous play, and hybrids like experience-weighted attaction learning (Camerer (2003)). BROAD can be applied to adaptively find games to distinguish solution concepts as well as different learning trajectories predicted by the theories.

Framing and Context Dependence The framing effect has been shown to be one of the most persistent biases of decision-making (Kahneman *et al.* (1982)). The framing of a task, and its context, affects the perception of the riskiness of choices (Kuhberger (1998) reports a meta-analysis), and the subjective probability that is assigned to outcomes (Rottenstreich and Hsee (2001)). Framing of the task also affects time perception. BROAD methods can be applied to compare theories in the relevant frame, or with the appropriate group of subjects in the field. Theories should also be compared in multiple frames or contexts to ensure the robustness of the results. By economizing the number of questions or tests required, BROAD can be used to test multiple frames in one, or just a few, experimental sessions.

5.1.3 Field experiments of Prospect Theory

The gold standard for inferring causal effects is controlled experiments. To test the causal effect of reference dependence and loss aversion on consumer behaviour, we would have to randomize the discounting periods and discount amounts for promotions of items and its substitute. In the case of the eCommerce retailer that we have studied, it is infeasbile to get the businesses' cooperation, since randomizing discounts could lead to losses or loss in consumer's confidence. The feasible alternative is to implement the pricing of an item and its substitute in order to maximize overall profits for both products, according to the prescriptions of our model. But now we randomly pick the pairs of items (treatment group) that implement this strategy, and use existing pricing criteria on the rest (control group). The observed change in demand between the two groups would test for the effectiveness of the strategy and strength of the causal effects.

Validating the effects in both lab and field studies addresses several issues of generalizability and control. An excellent recent example is Sonnemann *et al.* (2013), where evidence for partition dependence (which could lead to the sum of probabilities of outcomes exceeding 1) is obtained from prediction markets in the lab, 153 field studies of prediction markets and long-shot bets on horse races. Adaptive experimental design and the discrete choice framework provides the tools to make it easier on researchers to test behavioural theories and effects in different settings.

Bibliography

Abernethy, J., Evgeniou, T., Toubia, O., and Vert, J. (2007). Eliciting consumer preferences using robust adaptive choice questionnaires. *IEEE Transactions on Knowledge and Data Engineering*, pages 145–155.

Adams, J. S. (1963). Towards an understanding of inequity. *The Journal of Abnormal and Social Psychology*, **67**(5), 422.

Ainslie, G. and Haslam, N. (1992). Hyperbolic discounting. *Choice over time*, **5**, 57–92.

Alba, J., Mela, C., Shimp, T., and Urbany, J. (1999). The effect of discount frequency and depth on consumer price judgments. *Journal of Consumer Research*, **26**(2), 99–114.

Albrecht, K., Volz, K. G., Sutter, M., Laibson, D. I., and von Cramon, D. Y. (2011). What is for me is not for you: brain correlates of intertemporal choice for self and other. *Social cognitive and affective neuroscience*, **6**(2), 218–225.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine. *Econometrica: Journal of the Econometric Society*, pages 503–546.

Andreoni, J. and Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, **70**(2), 737–753.

Angrist, J. and Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. Technical report, National Bureau of Economic Research.

Arrow, K. (1971). *Essays in the theory of risk-bearing*. Markham Pub. Co.

Atkinson, A. and Donev, A. (1992). *Optimum Experimental Designs*. Oxford Univ Press.

Atkinson, A. and Fedorov, V. (1975). The design of experiments for discriminating between two rival models. *Biometrika*, **62**(1), 57.

Azfar, O. (1999). Rationalizing hyperbolic discounting. *Journal of Economic Behavior & Organization*, **38**(2), 245–252.

Balcan, N., Beygelzimer, A., and Langford, J. (2006). Agnostic active learning. In *International Conference in Machine Learning (ICML)*.

Barberis, N., Huang, M., and Santos, T. (2001). Prospect theory and asset prices*. *Quarterly Journal of Economics*, **116**(1), 1–53.

Bayroff, A., Thomas, J., and Anderson, A. (1960). Construction of an experimental sequential item test. Technical report, DTIC Document.

Bazerman, M. H., White, S. B., and Loewenstein, G. F. (1995). Perceptions of fairness in interpersonal and individual choice situations. *Current Directions in Psychological Science*, **4**(2), 39–43.

Becker, G., DeGroot, M., and Marschak, J. (1963). Stochastic models of choice behavior. *Behavioral Science*, **8**(1), 41–55.

Bell, D. and Lattin, J. (2000). Looking for loss aversion in scanner panel data: the confounding effect of price response heterogeneity. *Marketing Science*, pages 185–200.

Bell, D., Chiang, J., and Padmanabhan, V. (1999). The decomposition of promotional response: An empirical generalization. *Marketing Science*, pages 504–526.

Bellala, G., Bhavnani, S. K., and Scott, C. D. (2009). Group-based query learning for rapid diagnosis in time-critical situations. *CoRR*, **abs/0911.4511**.

Bellala, G., Bhavnani, S., and Scott, C. (2010). Extensions of generalized binary search to group identification and exponential costs. In *Advances in Neural Information Processing Systems (NIPS)*.

Benartzi, S. and Thaler, R. (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics*, **110**, 73–92.

Benhabib, J., Bisin, A., and Schotter, A. (2004). Hyperbolic discounting: An experimental analysis. Working paper, New York University Department of Economics.

Benhabib, J., Bisin, A., and Schotter, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior*, **69**(2), 205–223.

Bernheim, B. D. (2008). The psychology and neurobiology of judgment and decision making: Whats in it for economists? *Neuroeconomics: Decision making and the brain*, pages 115–26.

Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. *Proceedings of the St. Petersburg Imperial Academy of Sciences*, **5**(175–92).

Binet, A. and Simon, T. (1904). Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. *L'année Psychologique*, **11**(1), 191–244.

Binmore, K. G. (2009). *Rational Decisions*. Princeton University Press.

Bleichrodt, H., van Rijn, J., and Johannesson, M. (1999). Probability weighting and utility curvature in qaly-based decision making. *Journal of Mathematical Psychology*, **43**(2), 238–260.

Bleichrodt, H., Rohde, K. I., and Wakker, P. P. (2009). Non-hyperbolic time inconsistency. *Games and Economic Behavior*, **66**(11), 27–38.

Bolton, G. E. and Ockenfels, A. (2000). Erc: A theory of equity, reciprocity, and competition. *American economic review*, pages 166–193.

Bonetti, S. (1998). Experimental economics and deception. *Journal of Economic Psychology*, **19**(3), 377–395.

Booij, A., Van Praag, B., and Van de Kuilen, G. (2010). A parametric analysis of prospect theorys functionals for the general population. *Theory and Decision*, **68**(1), 115–148.

Briesch, R., Krishnamurthi, L., Mazumdar, T., and Raj, S. (1997). A comparative analysis of reference price models. *Journal of Consumer Research*, **24**(2), 202–214.

Buhusi, C. V. and Meck, W. H. (2005). What makes us tick? functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, **6**(10), 755–765.

Buonomano, D. V. (2007). The biology of time across different scales. *Nature chemical biology*, **3**(10), 594–597.

Burnashev, M. and Zigangirov, K. (1974). An interval estimation problem for controlled observations. *Problemy Peredachi Informatsii*, **10**(3), 51–61.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Camerer, C. (2005). Three cheers–psychological, theoretical, empirical–for loss aversion. *Journal of Marketing Research*, pages 129–133.

Camerer, C. (2011). The promise and success of lab-field generalizability in experimental economics: a critical reply to levitt and list. *SSRN Working Papers*.

Camerer, C., Babcock, L., Loewenstein, G., and Thaler, R. (1997). Labor supply of new york city cabdrivers: One day at a time. *The Quarterly Journal of Economics*, **112**(2), 407.

Camerer, C. F. (2004). Prospect theory in the wild: Evidence from the field. *Colin F. Camerer, George Loewenstein, and Matthew. Rabin, eds., Advances in Behavioral Economics*, pages 148–161.

Camerer, C. F. and Loewenstein, G. (2004). Behavioral economics: Past, present, future. *Advances in behavioral economics*, **3**.

Camerer, C. F., Ho, T.-H., and Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, **119**(3), 861–898.

Cavagnaro, D., Myung, J., Pitt, M., and Kujala, J. (2010). Adaptive design optimization: A mutual information-based approach to model discrimination in cognitive science. *Neural computation*, **22**(4), 887–905.

Cavagnaro, D., Pitt, M., and Myung, J. (2011). Model discrimination through adaptive experimentation. *Psychonomic bulletin & review*, **18**(1), 204–210.

Cavagnaro, D., Gonzalez, R., Myung, J., and Pitt, M. (Fortcoming). Optimal stimuli for decision making experiments: An adaptive approach. *Management Science*.

Chakaravarthy, V. T., Pandit, V., Roy, S., Awasthi, P., and Mohania, M. (2007). Decision trees for entity identification: Approximation algorithms and hardness results. In *In Proceedings of the ACM- SIGMOD Symposium on Principles of Database Systems*.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, **10**(3), 273–304.

Chamberlin, E. H. (1948). An experimental imperfect market. *The Journal of Political Economy*, **56**(2), 95–108.

Chambers, C. and Echenique, F. (2008). Ordinal notions of submodularity. *Journal of Mathematical Economics*, **44**(11), 1243–1245.

Chapman, C., Alford, J., Johnson, C., Lahav, M., and Weidemann, R. (2009). Comparing results of cbc and acbc with real product selection. In *Sawtooth Software Conference 2009*.

Chapman, G. B. and Elstein, A. S. (1995). Valuing the future temporal discounting of health and money. *Medical Decision Making*, **15**(4), 373–386.

Charness, G. and Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, **117**(3), 817–869.

Cheng, B. and Titterington, D. M. (1994). Neural networks: A review from a statistical perspective. *Statistical science*, pages 2–30.

Choi, S., Fisman, R., Gale, D., and Kariv, S. (2007). Consistency and heterogeneity of individual behavior under uncertainty. *The American economic review*, **97**(5), 1921–1938.

Christopoulos, G., Tobler, P., Bossaerts, P., Dolan, R., and Schultz, W. (2009). Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *The Journal of Neuroscience*, **29**(40), 12574.

Chung, S.-H. and Herrnstein, R. J. (1967). Choice and delay of reinforcement. *Journal of the Experimental Analysis of Behavior*, **10**(1), 67–74.

Crawford, V. P. and Meng, J. (2011). New york city cab drivers' labor supply revisited: Reference-dependent preferences with rationalexpectations targets for hours and income. *The American Economic Review*, **101**(5), 1912–1932.

Dasgupta, P. and Maskin, E. (2005). Uncertainty and hyperbolic discounting. *The American Economic Review*, **95**(4), 1290–1299.

Dasgupta, S. (2004). Analysis of a greedy active learning strategy. In *Advances in Neural Information Processing Systems (NIPS)*.

Dasgupta, S. (2006). Coarse sample complexity bounds for active learning. In *Advances in Neural Information Processing Systems (NIPS)*.

Daskalakis, C., Goldberg, P. W., and Papadimitriou, C. H. (2009). The complexity of computing a nash equilibrium. *SIAM Journal on Computing*, **39**(1), 195–259.

Davidson, D. and Marschak, J. (1959). Experimental tests of a stochastic decision theory. *Measurement: definitions and theories*, pages 233–69.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, **441**(7095), 876–879.

De Martino, B., Kumaran, D., Seymour, B., and Dolan, R. J. (2006a). Frames, biases, and rational decision-making in the human brain. *Science*, **313**(5787), 684–687.

De Martino, B., Kumaran, D., Seymour, B., and Dolan, R. (2006b). Frames, biases, and rational decision-making in the human brain. *Science*, **313**(5787), 684.

De Martino, B., Camerer, C. F., and Adolphs, R. (2010). Amygdala damage eliminates monetary loss aversion. *Proceedings of the National Academy of Sciences*, **107**(8), 3788–3792.

De Martino, B., O'Doherty, J. P., Ray, D., Bossaerts, P., and Camerer, C. (2013). In the mind of the market: theory of mind biases value computation during financial bubbles. *Neuron*, **forthcoming**.

DellaVigna, S. and Malmendier, U. (2006). Paying not to go to the gym. *The American Economic Review*, pages 694–719.

Ding, M. (2007). An incentive-aligned mechanism for conjoint analysis. *Journal of Marketing Research*, **44**(2), 214–223.

Ding, M., Grewal, R., and Liechty, J. (2005). Incentive-aligned conjoint analysis. *Journal of Marketing Research*, **67–82**.

D.J. Navarro, M. P. and Myung, I. (2004). Assessing the distinguishability of models and the informativeness of data. *Cognitive Psychology*, **49**, 47–84.

Dong, S., Ding, M., and Huber, J. (2010). A simple mechanism to incentive-align conjoint experiments. *International Journal of Research in Marketing*, **27**(1), 25–32.

Doyle, J. R. (2013). Survey of time preference, delay discounting models. *Judgment and Decision Making*, **8**(2), 116–135.

dAcremont, M. and Bossaerts, P. (2008). Neurobiological studies of risk assessment: a comparison of expected utility and mean-variance approaches. *Cognitive, Affective, & Behavioral Neuroscience*, **8**(4), 363–374.

Ebert, J. E. and Prelec, D. (2007). The fragility of time: Time-insensitivity and valuation of the near and far future. *Management Science*, **53**(9), 1423–1438.

Edgeworth, F. Y. (1881). *Mathematical psychics: An essay on the application of mathematics to the moral sciences*. Number 10. C. Kegan Paul & Company.

El-Gamal, M. and Palfrey, T. (1993). Economical experiments: Bayesian efficient experimental design. *International Journal of Game Theory*, **25**(4), 495–517.

El-Gamal, M., McKelvey, R., and Palfrey, T. (1993). A bayesian sequential experimental study of learning in games. *Journal of the American Statistical Association*, pages 428–435.

Engelmann, D. and Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *The American Economic Review*, **94**(4), 857–869.

Esary, J. D., Proschan, F., and Walkup, D. W. (1967). Association of random variables, with applications. *The Annals of Mathematical Statistics*, **38**(5), 1466–1474.

Estle, S. J., Green, L., Myerson, J., and Holt, D. D. (2007). Discounting of monetary and directly consumable rewards. *Psychological Science*, **18**(1), 58–63.

Farmer, J. and Geneakopolos, J. (2009). Hyperbolic discounting is rational: valuing the far future with uncertain discount rates. Discussion Paper 1719, Cowles Foundation.

Fechner, G. (1860). Elemente dur psychophysik. *Leipzig, Germany: Breitkopf and Hartel*.

Fehr, E. and Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in cognitive sciences*, **11**(10), 419–427.

Fehr, E. and Fischbacher, U. (2005). The economics of strong reciprocity. *Moral Sentiments and Material Interests. The Foundations for Cooperation in Economic Life*, pages 151–193.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, **114**(3), 817–868.

Fehr, E., Naef, M., and Schmidt, K. M. (2006). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Comment. *The American economic review*, **96**(5), 1912–1917.

Figner, B. and Murphy, R. O. (2010). Using skin conductance in judgment and decision making research. *A Handbook of Process Tracing Methods for Decision Research: A Critical Review and Users Guide*, pages 163–84.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, **10**(2), 171–178.

Fisman, R., Kariv, S., and Markovits, D. (2007). Individual preferences for giving. *The American Economic Review*, **97**(5), 1858–1876.

Fox, C. and Poldrack, R. (2009). Prospect theory and the brain. *Neuroeconomics. Decision Making and the Brain*, pages 145–173.

Frederick, S., Loewenstein, G., and O'donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of economic literature*, **40**(2), 351–401.

Friedman, M. (1953). *Essays in positive economics*, volume 231. University of Chicago Press.

Fujishige, S. (1978). Polymatroidal dependence structure of a set of random variables. *Information and Control*, **39**(1), 55–72.

Genesove, D. and Mayer, C. (2001). Loss aversion and seller behavior: Evidence from the housing market. *The Quarterly Journal of Economics*, **116**(4), 1233.

Gershon, R., Rothrock, N., Hanrahan, R., Bass, M., and Cella, D. (2010). The use of promis and assessment center to deliver patient-reported outcome measures in clinical research. *Journal of applied measurement*, **11**(3).

Gibbon, J., Church, R. M., and Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of sciences*, **423**(1), 52–77.

Gilboa, I., Postlewaite, A., and Schmeidler, D. (2009). The complexity of the consumer problem and mental accounting. Technical report, Pinhas Sapir Center for Development, Tel Aviv University.

Glimcher, P. W. *et al.* (2009). Choice: towards a standard back-pocket model. *Neuroeconomics: Decision making and the brain*, pages 503–521.

Gneezy, U. and List, J. (2006). Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, **74**(5), 1365–1384.

Golovin, D. and Krause, A. (2010a). Adaptive submodularity: A new approach to active learning and stochastic optimization. In *Proceedings of International Conference on Learning Theory (COLT)*. Citeseer.

Golovin, D. and Krause, A. (2010b). Adaptive submodularity: A new approach to active learning and stochastic optimization. In *Proc. of the 23rd Annual Conference on Learning Theory (COLT)*, pages 333–345.

Golovin, D. and Krause, A. (2010c). Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *CoRR*, **abs/1003.3967v3**.

Golovin, D. and Krause, A. (2011). Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research*, **42**, 427–486.

Golovin, D., Krause, A., and Ray, D. (2010). Near-optimal bayesian active learning with noisy observations. In *Advances in Neural Information Processing Systems (NIPS)*.

Greenleaf, E. (1995). The impact of reference price effects on the profitability of price promotions. *Marketing science*, pages 82–104.

Guillory, A. and Bilmes, J. (2009). Average-case active learning with costs. In *The 20th International Conference on Algorithmic Learning Theory*, University of Porto, Portugal.

Gul, F. and Pesendorfer, W. (2008). The case for mindless economics. *The foundations of positive and normative economics*, pages 3–39.

Gurevich, G., Kliger, D., and Levy, O. (2009). Decision-making under uncertainty-a field study of cumulative prospect theory. *Journal of Banking & Finance*, **33**(7), 1221–1229.

Hanoch, G. and Levy, H. (1969). The efficiency analysis of choices involving risk. *The Review of Economic Studies*, **36**(3), 335–346.

Hardisty, D. J., Appelt, K. C., and Weber, E. U. (2012). Good or bad, we want it now: Fixed-cost present bias for gains and losses explains magnitude asymmetries in intertemporal choice. *Journal of Behavioral Decision Making*.

Harless, D. and Camerer, C. (1994). The predictive utility of generalized expected utility theories. *Econometrica: Journal of the Econometric Society*, pages 1251–1289.

Harvey, C. M. (1986). Value functions for infinite-period planning. *Management Science*, **32**(9), 1123–1139.

Heath, C., Huddart, S., and Lang, M. (1999). Psychological factors and stock option exercise. *Quarterly Journal of Economics*, **114**(2), 601–627.

Hertwig, R. and Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences*, **24**(03), 383–403.

Hey, J. (2005). Why we should not be silent about noise. *Experimental Economics*, **8**(4), 325–345.

Hey, J. and Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica: Journal of the Econometric Society*, pages 1291–1326.

Ho, T. H., Lim, N., and Camerer, C. F. (2006). Modeling the psychology of consumer and firm behavior with behavioral economics. *Journal of Marketing Research*, pages 307–331.

Holt, C. and Laury, S. (2005). Risk aversion and incentive effects: New data without order effects. *The American economic review*, **95**(3), 902–904.

Howard, R. A. (1966). Information value theory. In *IEEE Transactions on Systems Science and Cybernetics (SSC-2)*.

Hsu, M., Krajbich, I., Zhao, C., and Camerer, C. (2009). Neural response to reward anticipation under risk is nonlinear in probabilities. *The Journal of Neuroscience*, **29**(7), 2231.

Hu, Y., Kayaba, Y., and Shum, M. (2013). Nonparametric learning rules from bandit experiments: the eyes have it! *Games and Economic Behavior*.

Jaynes, E. T. (2003). *Probability theory: the logic of science*. Cambridge university press.

Jevons, W. S. (1870). On the natural laws of muscular exertion. *Nature*, **2**, 158–60.

Johnson, R., Huber, J., and Bacon, L. (2003). Adaptive choice-based conjoint. In *Sawtooth Software Conference Proceedings.*

Kable, J. W. and Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, **10**(12), 1625–1633.

Kahneman, D. and Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive psychology*, **3**(3), 430–454.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47**(2), 263–292.

Kahneman, D., Slovic, P., and Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1990). Experimental tests of the endowment effect and the coase theorem. *Journal of political Economy*, pages 1325–1348.

Kalyanaram, G. and Winer, R. (1995). Empirical generalizations from reference price research. *Marketing Science*, pages 161–169.

Kang, M. J., Ray, D., and Camerer, C. (2012). Measured anxiety and choices in experimental timing games. Working paper, California Institute of Technology.

Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society Series B*, **21**(2), 272–319.

Kim, B. K. and Zauberman, G. (2009). Perception of anticipatory time in temporal discounting. *Journal of Neuroscience, Psychology, and Economics*, **2**(2), 91.

Kosaraju, S. R., Przytycka, T. M., and Borgstrom, R. S. (1999). On an optimal split tree problem. In *WADS '99: Proceedings of the 6th International Workshop on Algorithms and Data Structures*.

Koszegi, B. and Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, **121**(4), 1133–1165.

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, **13**(10), 1292–1298.

Krause, A. and Golovin, D. (2013). Submodular function maximization. *Tractability: Practical Approaches to Hard Problems*, **3**.

Krause, A. and Guestrin, C. (2009). Optimal value of information in graphical models. *Journal of Artificial Intelligence Research (JAIR)*, **35**, 557–591.

Kroll, Y., Levy, H., and Markowitz, H. (1984). Mean-variance versus direct utility maximization. *Journal of Finance*, pages 47–61.

Krugman, P. (2010). The theory of interstellar trade. *Economic Inquiry*, **48**(4), 1119–1123.

Kuhberger, A. (1998). The influence of framing on risky decisions: A meta-analysis. *Organizational Behavior and Human Decision Processes*, **75**(1), 23–55.

Kujala, J. and Lukka, T. (2006). Bayesian adaptive estimation: the next dimension. *Journal of Mathematical Psychology*, **50**(369-389).

L.A. Lesmes, S.-T. Jeon, Z.-L. L. and Dosher, B. A. (2006). Bayesian adaptive estimation of threshold versus contrast external noise functions: The quick tvc method. *Vision Research*, **46**(3160-3176).

Laibson, D. (1997). Golden eggs and hyperbolic discounting*. *Quarterly Journal of Economics*, **112**(2), 443–477.

Leskovec, J., Krause, A., Guestrin, C., Faloutsos, C., VanBriesen, J., and Glance, N. (2007). Cost-effective outbreak detection in networks. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*.

Levitt, S. D. and List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives*, **21**(2), 153–174.

Levy, H. and Markowitz, H. (1979). Approximating expected utility by a function of mean and variance. *The American Economic Review*, **69**(3), 308–317.

Lewi, J., Butera, R., and Paninski, L. (2009). Sequential optimal design of neurophysiology experiments. *Neural Computation*, **21**(3), 619–687.

Lindley, D. V. (1956). On a measure of the information provided by an experiment. *Annals of Mathematical Statistics*, **27**, 986–1005.

Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., and Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *The Journal of Neuroscience*, **21**(4), 1370–1377.

Linkenkaer-Hansen, K., Nikulin, V. V., Palva, J. M., Kaila, K., and Ilmoniemi, R. J. (2004). Stimulus-induced change in long-range temporal correlations and scaling behaviour of sensorimotor oscillations. *European Journal of Neuroscience*, **19**(1), 203–218.

Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, **107**(2), 573.

Loewenstein, G. F., Thompson, L., and Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social psychology*, **57**(3), 426.

Loomes, G. and Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, **39**(3), 641–648.

Loomes, G. and Sugden, R. (1998). Testing different stochastic specificationsof risky choice. *Economica*, **65**(260), 581–598.

Loomes, G., Moffatt, P., and Sugden, R. (2002). A microeconometric test of alternative stochastic theories of risky choice. *Journal of risk and Uncertainty*, **24**(2), 103–130.

Lord, F. (1974). Estimation of latent ability and item parameters when there are omitted responses. *Psychometrika*, **39**(2), 247–264.

Lovász, L. (1983). Submodular functions and convexity. In *Mathematical Programming The State of the Art*, pages 235–257. Springer.

Luce, R. and Krumhansl, C. (1988). Measurement, scaling, and psychophysics. *Stevens handbook of experimental psychology*, **1**, 3–74.

Machado, A. and Keen, R. (1999). Learning to time (let) or scalar expectancy theory (set)? a critical test of two models of timing. *Psychological Science*, **10**(3), 285–290.

MacKay, D. (1992). Information-based objective functions for active data selection. *Neural Computation*, **4**(4), 590–604.

Madden, G. J. and Bickel, W. K. (2010). Impulsivity: The behavioral and neurological science of discounting.

Mandelbrot, B. B. and Van Ness, J. W. (1968). Fractional brownian motions, fractional noises and applications. *SIAM review*, **10**(4), 422–437.

Markowitz, H. (1968). *Portfolio selection: Efficient diversification of investments*. Number 16. Yale Univ Pr.

Marschak, J. (1950). Rational behavior, uncertain prospects, and measurable utility. *Econometrica: Journal of the Econometric Society*, pages 111–141.

Mazumdar, T. and Papatla, P. (2000). An investigation of reference price segments. *Journal of Marketing Research*, pages 246–258.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement.

McClelland, G. (1997). Optimal design in psychological research. *Psychological Methods*, **2**(1).

McFadden, D. (1980). Econometric models for probabilistic choice among products. *Journal of Business*, pages 13–29.

McFadden, D. (2001). Economic choices. *The American Economic Review*, **91**(3), 351–378.

McFadden, D. and Train, K. (2000). Mixed mnl models for discrete response. *Journal of Applied Econometrics*, **15**(5), 447–470.

McKerchar, T. L., Green, L., Myerson, J., Pickford, T. S., Hill, J. C., and Stout, S. C. (2009). A comparison of four models of delay discounting in humans. *Behavioural Processes*, **81**(2), 256–259.

McKerchar, T. L., Green, L., and Myerson, J. (2010). On the scaling interpretation of exponents in hyperboloid models of delay and probability discounting. *Behavioural processes*, **84**(1), 440–444.

Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Cognitive Brain Research*, **3**(3), 227–242.

Meyer, R. and Johnson, E. (1995). Empirical generalizations in the modeling of consumer choice. *Marketing Science*, pages 180–189.

Minoux, M. (1978). Accelerated greedy algorithms for maximizing submodular set functions. In *Proc. of the 8th IFIP Conference on Optimization Techniques*, pages 234–243.

Mischel, W., Shoda, Y., and Rodriguez, M. L. (1992). Delay of gratification. *Choice over time*, **147**.

Moffatt, P. (2005). Stochastic choice and the allocation of cognitive effort. *Experimental Economics*, **8**(4), 369–388.

Moffatt, P. (2007). *Optimal experimental designs in models of decision and choice.*, chapter 15. Elsevier.

Moffatt, P. and Peters, S. (2001). Testing for the presence of a tremble in economic experiments. *Experimental Economics*, **4**(3), 221–228.

Monroe, K. (1973). Buyers' subjective perceptions of price. *Journal of Marketing Research*, pages 70–80.

Muller, P., Berry, D., Grieve, A., Smith, M., and Krams, M. (2007). Simulation-based sequential bayesian design. *Journal of Statistical Planning and Inference*, **137**, 3140–3150.

Myung, J. and Pitt, M. (2009). Optimal experimental design for model discrimination. *Psychological review*, **116**(3), 499.

Myung, J., Pitt, M., Tang, Y., and Cavagnaro, D. (2009). Bayesian adaptive optimal design of psychology experiments. In *Proceedings of the 2nd International Workshop in Sequential Methodologies (IWSM2009)*.

Nakahara, H. and Kaveri, S. (2010). Internal-time temporal difference model for neural value-based decision making. *Neural Computation*, **22**(12), 3062–3106.

Neal, R. M. (1995). *Bayesian learning for neural networks*. Ph.D. thesis, University of Toronto.

Nemhauser, G. L., Wolsey, L. A., and Fisher, M. L. (1978). An analysis of approximations for maximizing submodular set functions - I. *Mathematical Programming*, **14**(1), 265–294.

Nowak, R. (2009). Noisy generalized binary search. In *Advances in Neural Information Processing Systems (NIPS)*.

Odean, T. (1998). Are investors reluctant to realize their losses? *The Journal of finance*, **53**(5), 1775–1798.

Palacios-Huerta, I. and Serrano, R. (2006). Rejecting small gambles under expected utility. *Economics Letters*, **91**(2), 250–259.

Papadimitriou, C. (2003). *Computational complexity*. John Wiley and Sons Ltd.

Peirce, C. (1967). Note on the theory of economy of research. *Reprinted in Operations Research*, **15**(4), 643–648.

Phelps, E. S. and Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies*, **35**(2), 185–199.

Pine, S. and Weiss, D. (1978). The effects of knowledge of results and test difficulty on ability test performance and psychological reactions to testing. Psychometric methods program, research report 78-2., Minneapolis: University Minnesota, Department of Psychology.

Pine, S., Church, A., Gialluca, K., and Weiss, D. (1979). *Effects of computerized adaptive testing on Black and White students*. US Navy, Office of Naval Research.

Pinker, S. (1999). How the mind works. *Annals of the New York Academy of Sciences*, **882**(1), 119–127.

Pitt, L. D. (1982). Positively correlated normal variables are associated. *The Annals of Probability*, pages 496–499.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, **32**(1), 122–136.

Prelec, D. (1998). The probablity weighting function. *Econometrica*, **66**(3), 497–527.

Prelec, D. (2004). Decreasing impatience: A criterion for non-stationary time preference and hyperbolic discounting. *Scandinavian Journal of Economics*, **106**(3), 511–532.

Prevost, C., McNamee, D., Jessup, R., Bossaerts, P., and O'Doherty, J. P. (2013). Evidence for model-based computations in the human amygdala during pavlovian conditioning. *PLoS computational biology*, **9**(2).

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, **3**(4), 323–343.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American economic review*, pages 1281–1302.

Rachlin, H. (2006). Notes on discounting. *Journal of the experimental analysis of behavior*, **85**(3), 425–435.

Ray, D. and Bossaerts, P. (2011). Positive temporal dependence of the biological clock implies hyperbolic discounting. *Frontiers in neuroscience*, **5**.

Ray, D., Golovin, D., Krause, A., and Camerer, C. (2013). Bayesian rapid optimal adaptive design (broad): Method and application distinguishing models of risky choice. Working paper, California Institute of Technology.

Read, D., Frederick, S., Orsel, B., and Rahman, J. (2005). Four score and seven years from now: The date/delay effect in temporal discounting. *Management Science*, pages 1326–1335.

Reutskaja, E., Nagel, R., Camerer, C. F., and Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *The American Economic Review*, **101**(2), 900–926.

Rosato, A. (2013). Selling substitute goods to loss-averse consumers: Limited availability, bargains and rip-offs. Job market paper, University of California at Berkeley.

Roth, A. E. and Kagel, J. H. (1995). *The handbook of experimental economics*, volume 1. Princeton university press Princeton.

Rottenstreich, Y. and Hsee, C. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Science*, **12**(3), 185–190.

Ryan, B. (2007). *Corporate finance and valuation*. Cengage Learning EMEA.

Schoemaker, P. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature*, pages 529–563.

Scholten, M. and Read, D. (2010). The psychology of intertemporal tradeoffs. *Psychological Review*, **117**(3), 925.

Schrijver, A. (2003). *Combinatorial optimization : polyhedra and efficiency*. Volume B, Part IV, Chapters 39-49. Springer.

Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, **19**(3), 425–442.

Shirai, K. (2010). On the existence of a submodular utility function. Technical report, Mimeo: Waseda University.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, **69**(1), 99–118.

Simon, H. A. (1990). Invariants of human behavior. *Annual review of psychology*, **41**(1), 1–20.

Skiba, P. and Tobacman, J. (2008). Payday loans, uncertainty and discounting: explaining patterns of borrowing, repayment, and default. *Vanderbilt Law and Economics Research Paper*, (08-33).

Skog, O.-J. (1997). The strength of weak will. *Rationality and Society*, **9**(2), 245–271.

Smith, P. L. and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in neurosciences*, **27**(3), 161–168.

Smith, V. L. (1962). An experimental study of competitive market behavior. *The Journal of Political Economy*, **70**(2), 111–137.

Smith, V. L. (1982). Microeconomic systems as an experimental science. *The American Economic Review*, **72**(5), 923–955.

Smith, V. L., Suchanek, G. L., and Williams, A. W. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica: Journal of the Econometric Society*, pages 1119–1151.

Sonnemann, U., Camerer, C. F., Fox, C. R., and Langer, T. (2013). How psychological framing affects economic market prices in the lab and field. *Proceedings of the National Academy of Sciences*, **110**(29), 11779–11784.

Sozou, P. D. (1998). On hyperbolic discounting and uncertain hazard rates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **265**(1409), 2015–2020.

Sozou, P. D. and Seymour, R. M. (2003). Augmented discounting: interaction between ageing and time–preference behaviour. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**(1519), 1047–1053.

Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. (2012). Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *Information Theory, IEEE Transactions on*, **58**(5), 3250–3265.

Strotz, R. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, **23**(3), 165–180.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement learning: An introduction*, volume 1. Cambridge Univ Press.

Swartz, R. and Choi, S. (2009). A burdened cat: Incorporating response burden with maximum fisher's information for item selection. In *Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing*.

Symmonds, M., Bossaerts, P., and Dolan, R. J. (2010). A behavioral and neural evaluation of prospective decision-making under risk. *The Journal of Neuroscience*, **30**(43).

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, **1**(1), 39–60.

Thurstone, L. L. (1931). The indifference function. *The Journal of Social Psychology*, **2**(2), 139–167.

Tobin, J. (1958). Liquidity preference as behavior towards risk. *The Review of Economic Studies*, **25**(2), 65–86.

Tom, S., Fox, C., Trepel, C., and Poldrack, R. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, **315**(5811), 515.

Toubia, O., Hauser, J., and Simester, D. (2004). Polyhedral methods for adaptive choice-based conjoint analysis. *Journal of Marketing Research*, pages 116–131.

Toubia, O., Johnson, E., Evgeniou, T., and Delquie, P. (Forthcoming). Dynamic experiments for estimating preferences: An adaptive method of eliciting time and risk parameters. *Management Science*.

Train, K. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Trepel, C., Fox, C., and Poldrack, R. (2005). Prospect theory on the brain? toward a cognitive neuroscience of decision under risk. *Cognitive Brain Research*, **23**(1), 34–50.

Tsukayama, E. and Duckworth, A. L. (2010). Domain-specific temporal discounting and temptation. *Judgment and Decision Making*, **5**(2), 72–82.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, **5**(4), 297–323.

Van Der Linden, W. and Glas, C. (2000). *Computerized adaptive testing: Theory and practice*. Springer.

Von Gaudecker, H.-M., Van Soest, A., and Wengstrom, E. (2011). Heterogeneity in risky choice behavior in a broad population. *The American Economic Review*, **101**(2), 664–694.

von Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behaviour*. Princeton University Press.

Wainer, H. (1990). *Computerized adaptive testing*. Wiley Online Library.

Wakker, P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge Univ Pr.

Wald, A. (1949). Statistical decision functions. *The Annals of Mathematical Statistics*, pages 165–205.

Wang, J. T.-y., Spezio, M., and Camerer, C. F. (2010a). Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *The American Economic Review*, pages 984–1007.

Wang, S., Filiba, M., and Camerer, C. F. (2010b). Dynamically optimized sequential experimentation (dose) for estimating economic preference parameters. *Under Review*.

Weatherly, J. N., Terrell, H. K., and Derenne, A. (2010). Delay discounting of different commodities. *The Journal of General Psychology: Experimental, Psychological, and Comparative Psychology*, **137**(3), 273–286.

Weiss, D. (1975). Computerized adaptive ability measurement. Technical report, Naval Research Reviews.

Weiss, D. and Betz, N. (1973). *Ability measurement: Conventional or adaptive?* Psychometric Methods Program, Department of Psychology, University of Minnesota.

Wilcox, N. (2008). Predicting individual risky choices out-of-context: a critical stochastic modeling primer and monte carlo study. *Risk Aversion in Experiments*.

Wittmann, M. and Paulus, M. P. (2008). Decision-making, impulsivity and time perception. *Trends in Cognitive Science*, **12**(1), 7–12.

Wittmann, M., Leland, D. S., Churan, J., and Paulus, M. P. (2007). Impaired time perception and motor timing in stimulant-dependent subjects. *Drug and alcohol dependence*, **90**(2), 183–192.

Wolsey, L. A. (1982). An analysis of the greedy algorithm for the submodular set covering problem. *Combinatorica*, **2**(4), 385–393.

Xiang, T., Ray, D., Lohrenz, T., Dayan, P., and Montague, P. R. (2012). Computational phenotyping of two-person interactions reveals differential neural response to depth-of-thought. *PLoS computational biology*, **8**(12), e1002841.

Zauberman, G., Kim, B. K., Malkoc, S. A., and Bettman, J. R. (2009). Discounting time and time discounting: Subjective time perception and intertemporal preferences. *Journal of Marketing Research*, **46**(4), 543–556.

Zheng, A. X., Rish, I., and Beygelzimer, A. (2005). Efficient test selection in active diagnosis via entropy approximation. In *UAI '05, Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence*.